

IMPLEMENTASI ALGORITME K-MEANS CLUSTERING DAN NAIVE BAYES CLASSIFIER UNTUK KLASIFIKASI DIAGNOSA PENYAKIT PADA KUCING

SKRIPSI

Untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun oleh:
Puji Indah Lestari
NIM: 12515010711010



**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BRAWIJAYA
MALANG
2018**

PENGESAHAN

IMPLEMENTASI ALGORITME K-MEANS CLUSTERING DAN NAIVE BAYES CLASSIFIER UNTUK KLASIFIKASI DIAGNOSA PENYAKIT PADA KUCING

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun Oleh :
Puji Indah Lestari
1251501071111010

Skripsi ini telah diuji dan dinyatakan lulus pada
3 Agustus 2018

Telah diperiksa dan disetujui oleh:

Dosen Pembimbing I

Dosen Pembimbing II

Dian Eka Ratnawati, S.Si., M.Kom
NIK: 19730619 200212 2 001

Lailil Muflikhah, S.Kom., M.Kom
NIK: 19741113 200501 2 001

Mengetahui
Ketua Jurusan Teknik Informatika

Tri Astoto Kurniawan, S.T., M.T., Ph.D
NIP: 19710518 200312 1 001

PERNYATAAN ORISINALITAS

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar pustaka.

Apabila ternyata didalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur plagiasi, saya bersedia skripsi ini digugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70).

Malang, 06 Agustus 2018

Puji Indah Lestari

125150107111010



KATA PENGANTAR

Dengan menyebut nama Allah SWT Yang Maha Pengasih dan Maha Penyayang. Puji syukur kehadiran Allah SWT karena limpahan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING DAN NAIVE BAYES CLASSIFIER UNTUK KLASIFIKASI DIAGNOSA PENYAKIT PADA KUCING ”. Shalawat serta salam senantiasa tercurahkan kepada junjungan Nabi besar kita Nabi Muhammad SAW beserta keluarga dan para sahabatnya. Skripsi ini disusun untuk memenuhi salah satu syarat memperoleh gelar Sarjana Komputer di Fakultas Ilmu Komputer di Universitas Brawijaya Malang.

Pada kesempatan ini penulis juga ingin menyampaikan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan bantuan, inspirasi kebaikan dan dukungan dalam menyelesaikan skripsi ini, terutama kepada:

1. Kedua Orang Tua Penulis yaitu Ayah, Ibu dan kedua saudara penulis yaitu Utami dan Mita, yang selalu memberi do’a, dukungan dan motivasi dengan tulus kepada penulis.
2. Dian Eka Ratnawati, S.Si., M.Kom., selaku dosen pembimbing utama yang telah meluangkan waktu, membimbing, mengarahkan dan memberikan saran kepada penulis sehingga dapat menyelesaikan skripsi ini.
3. Lailil Munziah, S.Kom., M.Kom., selaku dosen pembimbing kedua yang telah meluangkan waktu, membimbing, mengarahkan dan memberikan saran kepada penulis sehingga dapat menyelesaikan skripsi ini.
4. Drs. Achmad Ridok, M.Kom., selaku dosen pembimbing akademik.
5. Tri Astoto Kurniawan, S.T, M.T, Ph.D., selaku ketua jurusan Teknik Informatika Universitas Brawijaya.
6. Agus Wahyu Widodo, S.T., M.Sc., selaku Ketua Program Studi Teknik Informatika Universitas Brawijaya

7. Wayan Firdaus Mahmudy, S.Si., M.T., Ph.D, selaku ketua Fakultas Ilmu Komputer Universitas Brawijaya.
8. Seluruh Dosen FILKOM UB yang telah membagikan ilmunya dan menginspirasi dalam kebaikan kepada penulis selama masa perkuliahan.
9. Seluruh staff dan karyawan FILKOM UB yang telah banyak membantu dalam hal administrasi penulis dalam pelaksanaan penyusunan skripsi ini.
10. Keluarga penulis yaitu Nenek, paklek Longgar, bulek Tatik, om Woyo, tante Sut, sepupu Intan, Ayu, Akris, Heru, adik Kiky, Nanda, Rista, Ferdi, Rahma dan keluarga besar atas segala do'a, nasihat, dukungan baik moril maupun materil yang begitu besar terhadap kelancaran dalam menyelesaikan skripsi ini.
11. Sahabat serta teman-teman penulis, Zurya, Reza, Iwa, Yubi, Citra, Ayu, Tanti, Nining, Ossy, Zila, Eza, Uca, Indah, Nanda, Mb Risa, Meyda, serta seluruh teman-teman angkatan 2012, angkatan 2013, dan angkatan 2014 yang selalu memberikan bantuan, dukungan, motivasi dan berbagi informasi demi kelancaran skripsi.
12. Sahabat serta teman-teman penulis di Malang, Lampung dan Ponorogo yang selalu mendukung dan memotivasi demi kelancaran skripsi.

Dengan kerendahan hati, penulis menyadari bahwa skripsi ini masih memiliki banyak kekurangan. Oleh karena itu kritik dan saran yang membangun sangat dibutuhkan sebagai pedoman untuk menyempurnakan skripsi ini agar lebih baik. Penulis berharap semoga skripsi ini dapat bermanfaat bagi diri sendiri maupun bagi semua pihak.

Malang, 04 Juni 2018

Penulis

indahifub@gmail.com

ABSTRAK

Kucing dewasa ini telah menjadi hewan peliharaan yang digemari masyarakat. Hal tersebut dikarenakan banyak manfaat yang didapat dari memelihara kucing, diantaranya adalah sebagai penghibur, bahkan sekarang di negara maju kucing banyak dipertandingkan dalam ajang *cat show*. Perawatan terhadap kucing terutama masalah kesehatan sangatlah penting karena beberapa penyakit pada kucing dapat menular pada manusia. Keterbatasan pemilik hewan dalam mendiagnosis gejala penyakit dengan pola memiliki gejala yang sama diantara beberapa penyakit, menjadi penyebab pemilik peliharaan mengalami kesulitan dalam melakukan diagnosis. Oleh sebab itu perlu ada sebuah sistem yang dapat mendiagnosa penyakit sehingga pemilik peliharaan dapat segera mengetahui penyakitnya dan dapat segera ditangani. Pada penelitian ini dapat memanfaatkan metode *K-Means Naïve Bayes* (KMNB) untuk diagnosa pada kucing. Pendekatan KMNB terbentuk dengan penggabungan teknik *clustering* dan klasifikasi. Pada tahap awal *Clustering* pada *K-Means* dimanfaatkan untuk mengelompokkan data yang sama. Selanjutnya klasifikasi data berdasarkan kategori menggunakan metode *Naïve Bayes*. Data yang memiliki kesalahan pada tahap pertama kemudian diklasifikasikan berdasarkan kategori pada tahap yang kedua. Berdasar pada hasil pengujian yang telah dilakukan dengan membandingkan hasil dari pengelompokkan pada *K-Means* konvensional membuktikan bahwa KMNB dapat memberikan rata-rata akurasi tertinggi sebesar 90% sedangkan *K-Means* konvensional memiliki rata-rata akurasi tertinggi sebesar 71,379%.

Kata Kunci: penyakit kucing, pengelompokkan, klasifikasi, *K-Means*, *Naïve Bayes*, KMNB.

ABSTRACT

At this time cat has become a popular pet community. This is because there are many benefits that exist from cats, such as an entertainer, and now developed countries many cats contested in the show cat. Treatment for cats is mainly because some of the cat's disease can spread to humans. The limitations of dentists in diagnosing diseases with a pattern of having the same symptoms as some diseases, are important in making a diagnosis. Therefore there need a system that can diagnose diseases that can be accessed by the cat owners and be dealt immediately. In this study can use K-Means Naïve Bayes (KMNB) method for diagnosis in cats. The KMNB approach is formed by the incorporation of clustering and classification techniques. In the beginning Clustering on K-Means was used to group the same data. Further classification of data by category using Naïve Bayes method. The data that have errors in the first stage are then organized by the second category. Identify data with the same character or data that shows similar characteristics from the start. Based on the results of tests that have been done by comparing the results of grouping on conventional K-Means proves that KMNB can produce the highest average of 90% while conventional K-Means has the highest average of 71, 379%.

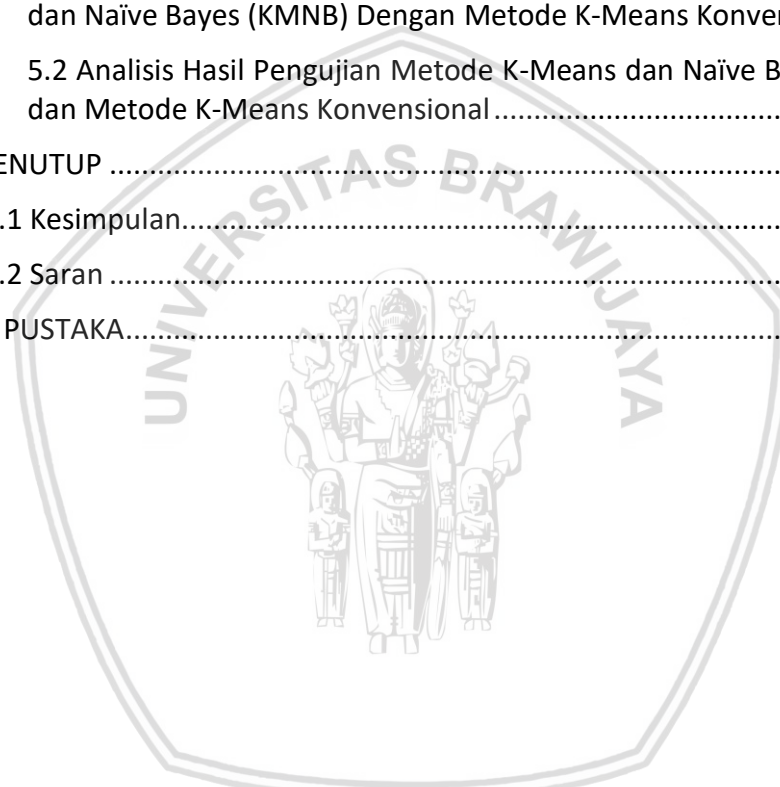
Keywords: cat disease, clustering, classification, K-Means, Naïve Bayes, KMNB.

DAFTAR ISI

PENGESAHAN	ii
PERNYATAAN ORISINALITAS	ii
KATA PENGANTAR.....	iiv
ABSTRAK.....	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR <i>SOURCE CODE</i>	xiii
BAB 1 PENDAHULUAN.....	1
1.1 Latar belakang.....	1
1.2 Rumusan masalah.....	3
1.3 Tujuan	3
1.4 Manfaat.....	4
1.5 Batasan masalah	4
1.6 Sistematika pembahasan.....	5
BAB 2 LANDASAN KEPUSTAKAAN	6
2.1 Kajian Pustaka	6
2.2 Klasifikasi.....	7
2.3 K-Means Clustering dan Naïve Bayes Classifier	8
2.3.1 K-Means Clustering	8
2.3.2 Naïve Bayes Classifier	9
2.4 KMNB (K-Means Naïve Bayes)	12
2.5 Penyakit Kucing.....	13
2.5.1 Scabies.....	13
2.5.2 Gastritis.....	13
2.5.3 Helminthiasis.....	14
2.5.4 Rhinitis.....	14
2.5.5 Dermatitis.....	15
2.5.6 Dermaphytosis	15

2.5.7 Otitis	16
2.5.8 Enteritis	17
BAB 3 METODOLOGI	18
3.1 Jenis Penelitian	18
3.2 Metodologi Penelitian	18
Langkah-langkah yang akan dilakukan dalam penelitian ini dapat di lihat pada Gambar 3.1.	18
3.2.1 Studi Literatur	19
3.2.2 Analisis Kebutuhan	19
3.2.3 Pengumpulan Data	19
3.2.4 Perancangan Sistem	19
3.2.5 Implementasi Sistem	47
3.2.6 Pengujian Sistem	48
3.2.7 Kesimpulan	49
BAB 4 IMPLEMENTASI DAN PEMBAHASAN	50
4.1 Lingkungan Implementasi	50
4.1.1 Lingkungan Implementasi Perangkat Keras	50
4.1.2 Lingkungan Implementasi Perangkat Lunak	51
4.2 Implementasi Program	51
4.2.1 Proses membuat <i>centroid</i> baru secara acak pada <i>K-Means</i>	51
4.2.2 Proses menghitung jarak data ke <i>centroid</i> pada <i>K-Means</i>	52
Source Code 4.2 Proses menghitung jarak data ke <i>centroid</i> pada <i>K-Means</i>	53
4.2.3 Proses meng- <i>update</i> (memperbaharui) <i>centroid</i>	53
Source Code 4.3 Proses meng- <i>update centroid</i>	54
4.2.4 Proses mendapatkan data latih untuk Naïve Bayes	55
Source Code 4.4 Proses mendapatkan data latih untuk <i>Naïve Bayes</i> .	55
4.2.5 Proses Menghitung <i>Prior Probability</i> pada Naïve Bayes	55
Source Code 4.5 Proses Menghitung <i>Prior Probability</i>	56
4.2.6 Proses perhitungan <i>Likelihood</i> dan <i>Posterior</i>	56
4.2.7 Proses Menghitung Akurasi	58
4.2.8 Proses menentukan klas hasil uji	59

Source Code 4.8 Proses Menentukan klas hasil uji.....	59
4.3 Implementasi Antarmuka	60
4.3.1 Antarmuka Utama	60
4.3.2 Antarmuka Aplikasi	61
BAB 5 PENGUJIAN DAN ANALISIS.....	63
5.1 Pengujian Sistem.....	63
5.1.1 Penentuan Data Latih dan Data Uji.....	63
5.1.2 Perbandingan Hasil Akurasi Antara Metode Gabungan K-Means dan Naïve Bayes (KMNB) Dengan Metode K-Means Konvensional.....	63
5.2 Analisis Hasil Pengujian Metode K-Means dan Naïve Bayes (KMNB) dan Metode K-Means Konvensional	66
BAB 6 PENUTUP	68
6.1 Kesimpulan.....	68
6.2 Saran	68
DAFTAR PUSTAKA.....	70

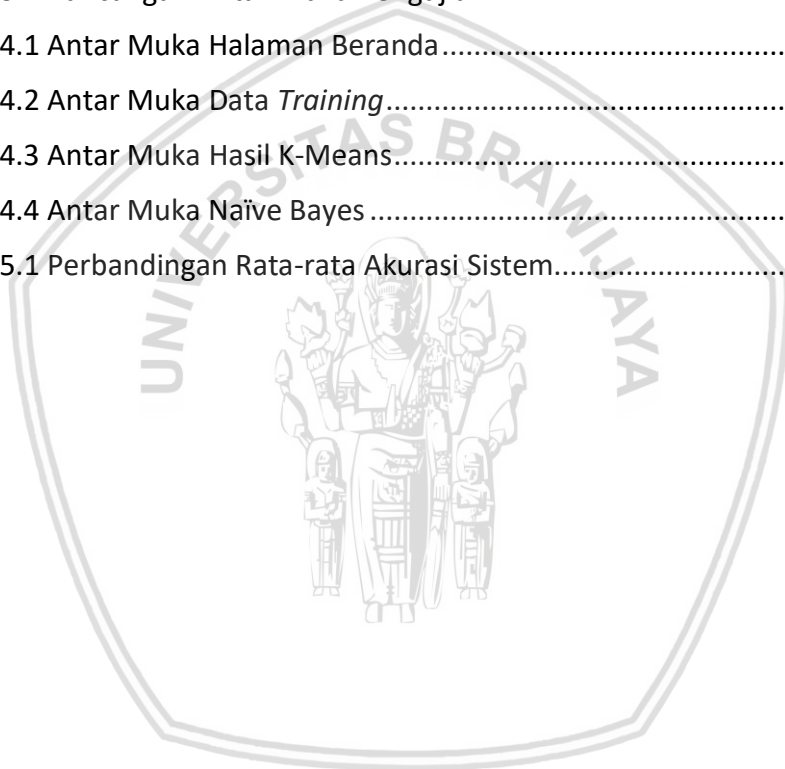


DAFTAR TABEL

Tabel 3.1 Data Gejala Penyakit Kucing.....	26
Tabel 3.2 Kode Gejala.....	28
Tabel 3.3 Pusat Kluster Awal	29
Tabel 3.4 Hasil Perhitungan Jarak <i>Eucliden</i> Iterasi ke-1.....	31
Tabel 3.5 Perhitungan Kluster Iterasi ke-1	33
Tabel 3.6 Hasil Perhitungan Jarak <i>Eucliden</i> Iterasi ke-2.....	34
Tabel 3.7 Hasil Perhitungan Jarak <i>Eucliden</i> Iterasi ke-3.....	36
Tabel 3.8 Pengelompokan Kluster.....	38
Tabel 3.9 Data C9	39
Tabel 3.10 Hasil Perhitungan <i>Prior</i> Tiap Kelas	39
Tabel 3.11 Hasil Perhitungan <i>likelihood</i> data ke-3.....	40
Tabel 3.12 Hasil Perhitungan <i>posterior</i> data ke-3.....	41
Tabel 3.13 Hasil Perhitungan <i>likelihood</i> data ke-40.....	42
Tabel 3.14 Hasil Perhitungan <i>posterior</i> data ke-40.....	43
Tabel 3.15 Hasil Pengujian	44
Tabel 3.16 Perancangan Uji Coba Sistem.....	48
Tabel 5.1 Akurasi K-Means Konvensional dengan Data Latih 80%.....	64
Tabel 5.2 Akurasi K-Means Konvensional dengan Data Latih 70%.....	64
Tabel 5.3 Akurasi K-Means Konvensional dengan Data Latih 60%.....	64
Tabel 5.4 Akurasi <i>K-Means Naïve Bayes</i> dengan Data Latih 80%.....	65
Tabel 5.5 Akurasi <i>K-Means Naïve Bayes</i> dengan Data Latih 70%.....	65
Tabel 5.6 Akurasi <i>K-Means Naïve Bayes</i> dengan Data Latih 60%.....	65
Tabel 5.7 Hasil Pengujian K-Means dan Naïve Bayes (KMNB) dan K-Means Konvensional	66

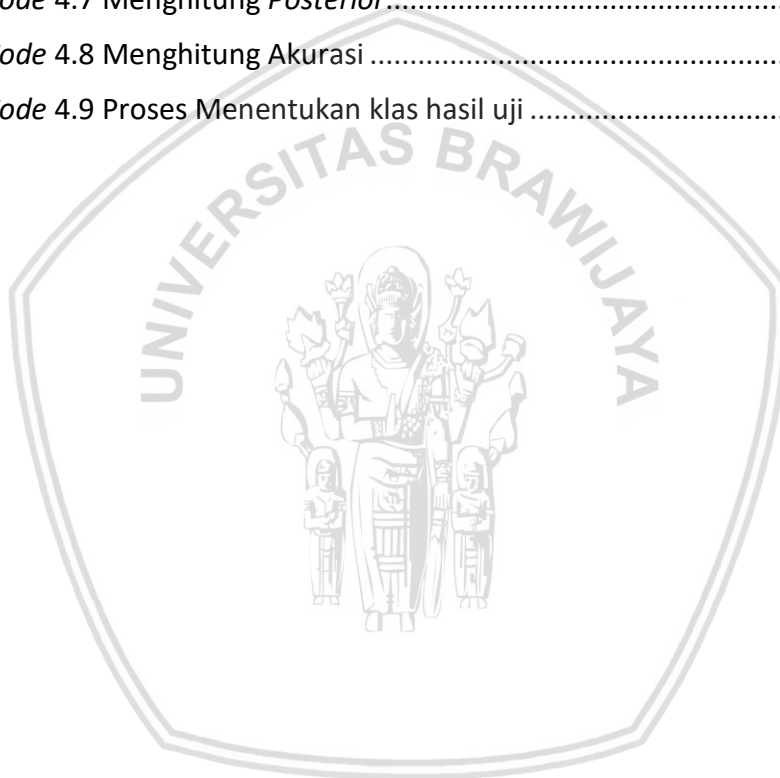
DAFTAR GAMBAR

Gambar 3.1 Diagram Metodologi Penelitian	18
Gambar 3.2 Alur Proses Sistem.....	21
Gambar 3.3 Alur Sistem Pengelompokan <i>K-Means</i>	22
Gambar 3.4 Alur Klasifikasi <i>Naïve Bayes</i>	24
Gambar 3.5 Rancangan Antar Muka Halaman Beranda	45
Gambar 3.6 Rancangan Antar Muka Data Training	46
Gambar 3.7 Rancangan Antar Muka Pengujian	47
Gambar 4.1 Antar Muka Halaman Beranda	60
Gambar 4.2 Antar Muka Data <i>Training</i>	61
Gambar 4.3 Antar Muka Hasil <i>K-Means</i>	62
Gambar 4.4 Antar Muka <i>Naïve Bayes</i>	62
Gambar 5.1 Perbandingan Rata-rata Akurasi Sistem.....	67



DAFTAR SOURCE CODE

Source Code 4.1 Proses membuat <i>centroid</i> baru secara acak pada <i>K-Means</i>	53
Source Code 4.2 Proses menghitung jarak data ke <i>centroid</i> pada <i>K-Means</i>	54
Source Code 4.3 Proses meng- <i>update centroid</i>	55
Source Code 4.4 Proses mendapatkan data latih untuk Naïve Bayes.....	56
Source Code 4.5 Proses Menghitung <i>Prior Probability</i>	56
Source Code 4.6 Proses menghitung <i>likelihood</i> pada <i>Naïve Bayes</i>	58
Source Code 4.7 Menghitung <i>Posterior</i>	59
Source Code 4.8 Menghitung Akurasi	59
Source Code 4.9 Proses Menentukan klas hasil uji	60



BAB 1 PENDAHULUAN

1.1 Latar belakang

Memelihara binatang peliharaan sekarang telah menjadi gaya hidup. Jenis hewan peliharaan yang paling sering dipelihara yaitu kucing. Banyak masyarakat di dunia yang memelihara kucing, termasuk masyarakat Indonesia (Susanti, 2005). Kucing merupakan hewan yang memiliki banyak kelebihan. Sifatnya yang manja membuat pecinta kucing selalu ingin mengelus dan bermain dengannya. Perilaku kucing yang lucu membuat pecintanya terhibur. Kucing berasal dari alam liar lalu perlahan-lahan mengalami proses domestikasi sehingga kini menjadi hewan peliharaan yang sangat dekat dengan manusia. Bahkan diberbagai negara maju kucing sudah banyak dipertandingkan dalam ajang *cat show* (Suwed dkk, 2011).

Perawatan terhadap kucing merupakan hal yang penting terutama masalah kesehatan. Ada beberapa penyakit pada kucing yang dapat menular pada manusia. Keterbatasan pemilik hewan dalam mendiagnosis gejala penyakit dengan pola memiliki gejala yang sama diantara beberapa penyakit, menjadi penyebab pemilik peliharaan mengalami kesulitan dalam melakukan diagnosis. Oleh karena itu perlu adanya sistem yang dapat mendiagnosa penyakit sehingga pemilik peliharaan dapat segera mengetahui penyakitnya dan dapat segera ditangani. Dengan adanya sistem klasifikasi dapat membantu mengolah data gejala penyakit dan memberikan keluaran sebagai informasi terkait diagnosa penyakit kucing.

Pada penelitian ini dapat memanfaatkan metode *K-Means Clustering* dan *Naïve Bayes Classifier* untuk diagnosa pada kucing. Sebagai referensi untuk membantu dalam pengerjaan skripsi ini, sebelumnya telah ada penelitian terkait metode *K-Means Clustering* dan *Naïve Bayes Classifier* dengan judul *A K-Means and Naïve Bayes Learning Approach For Better Intrusion Detection* yang dilakukan oleh Z. Muda dkk. Dalam penelitian ini pendekatan K-Means dan Naïve Bayes Classifier (KMNB) dilakukan untuk melakukan deteksi intrusi. Hasil

penelitian menunjukkan bahwa KMNB secara signifikan meningkatkan akurasi, tingkat deteksi, dan *false alarm* sampai 99.6, 99.8 dan 0,5% (Z. Muda dkk, 2011). Penelitian lainnya dengan judul *K-Means Clustering and Naïve Bayes Classifier For Categorization Of Diabetes Patients*. Penelitian tersebut untuk kategorisasi penderita diabetes yang dilakukan oleh Pandeewari dkk. Dalam penelitian tersebut menyimpulkan integrasi *clustering* dan teknik klasifikasi membantu dalam mengidentifikasi data set yang besar. Teknik ini memberikan hasil yang lebih akurat daripada teknik klasifikasi sederhana (Pandeewari dkk, 2015). Penelitian selanjutnya dengan judul *A Modified K-Means with Naïve Bayes (KMNB) Algorithm for Breast Cancer Classification* oleh Dian Eka Ratnawati dkk. Penelitian ini mengusulkan bahwa metode K-Means Naive Bayes (KMNB) dapat meningkatkan hasil klasifikasi pada data kanker payudara. Dengan melakukan modifikasi pada centroid awal dengan Fang. Keakuratan metode yang diusulkan adalah 95%, dan dapat mengurangi kesalahan sekitar 50%. Penelitian ini menyatakan bahwa metode KMNB cukup menjanjikan dan mampu meningkatkan prediksi penyakit kanker payudara (Ratnawati dkk, 2018).

Penelitian selanjutnya dilakukan oleh Anggreyni adalah *Implementasi Metode Clustering Untuk Klasifikasi Kanker Payudara Menggunakan Algoritme Naïve Bayes dan K-Means*. Pada penelitian ini digunakan *clustering* K-Means dan Naïve Bayes. Pada K-Means dilakukan pengelompokan data. Kemudian hasil 'mungkin' dicari peluangnya dengan menggunakan Naïve Bayes. Penelitian ini membuktikan penggunaan K-Means dan Naïve Bayes pada dataset kanker payudara memberikan akurasi tertinggi sebesar 97,72% dan akurasi terendah 96,02% (Anggreyni, 2015).

Penelitian terkait lainnya adalah *Penentuan Penerima Bantuan Ternak Menggunakan Algoritme K-Means & Naïve Bayes*, dilakukan oleh Mohamad Fadel Asikin. Penggabungan teknik klastering dan klasifikasi pada penelitian ini dianggap mampu memberikan hasil klasifikasi yang akurat pada penerimaan bantuan ternak. Pengelompokan data yang sama awal tahap dilakukan dengan pendekatan K-Means sebagai komponen pra-klasifikasi. Pendekatan selanjutnya

menggunakan Naïve Bayes untuk klasifikasi data berdasarkan kategori diterima atau tidaknya bantuan ternak. Berdasarkan hasil pengujian dari penelitian ini, hasil pengelompokkan menggunakan metode KMNB terbukti memberikan akurasi tertinggi sebesar 100% sedangkan K-Means konvensional memiliki akurasi sebesar 95.91 (Asikin, 2017).

Pendekatan KMNB dibentuk dengan menggabungkan teknik *clustering* dan klasifikasi. Teknik *clustering* pada K-Means digunakan untuk mengelompokkan data yang sama. Selanjutnya klasifikasi data berdasarkan kategori menggunakan metode *Naïve Bayes*. Dengan demikian data yang memiliki kesalahan pada tahap pertama akan diklasifikasi berdasarkan kategori pada tahap kedua. *Clustering* memberikan keuntungan yang signifikan atas teknik klasifikasi yang membantu mengidentifikasi kelompok data yang berkarakter sama atau menunjukkan karakteristik yang serupa di awal (Z. Muda, 2011).

Berdasarkan pemaparan diatas, maka akan dicoba pendekatan pembelajaran pada skripsi ini, yaitu *K-Means Clustering* dan *Naïve Bayes* (KMNB) untuk klasifikasi pada penyakit kucing dengan judul, "*Implementasi Metode K-Means Clustering dan Naïve Bayes Classifier untuk Diagnosa Penyakit Kucing*".

1.2 Rumusan masalah

Berdasarkan uraian diatas, maka dapat dirumuskan permasalahan pada penelitian ini yaitu sebagai berikut:

1. Bagaimana mengimplementasi algoritma *K-Means Clustering dan Naïve Bayes Classifier* untuk klasifikasi gejala diagnosa penyakit pada kucing?
2. Bagaimana tingkat akurasi yang diperoleh dari sistem klasifikasi untuk mendiagnosa penyakit pada kucing dengan metode *K-Means Clustering dan Naïve Bayes Classifier*?

1.3 Tujuan

Tujuan yang ingin dicapai dalam pembuatan skripsi ini adalah sebagai berikut:

1. Mengimplementasikan sistem klasifikasi untuk mendiagnosa penyakit pada kucing dengan metode *K-Means Clustering* dan *Naive Bayes Classifier*.
2. Menghitung tingkat akurasi yang diperoleh dari sistem klasifikasi untuk mendiagnosa penyakit pada kucing dengan metode *K-Means Clustering* dan *Naive Bayes Classifier*.

1.4 Manfaat

Penulisan penelitian ini diharapkan mempunyai manfaat yang baik dan berguna bagi pembaca dan penulis. Adapun manfaat yang diharapkan adalah dapat dijadikan referensi mengenai diagnosa penyakit pada kucing dan dapat memberikan kemudahan bagi pengguna dalam mendiagnosa penyakit dan menentukan penanggulangan penyakit pada kucing.

1.5 Batasan masalah

Untuk menghindari semakin melebarnya masalah, maka dari rumusan masalah yang telah dipaparkan, batasan masalah yang diterapkan adalah sebagai berikut:

1. Pengembangan implementasi untuk mendiagnosa penyakit pada kucing menggunakan metode *K-Means Clustering* dan *Naive Bayes Classifier*.
2. Sumber data berasal dari penelitian sebelumnya dimana data didapat dari buku dan data rekam medis pada klinik Universitas Brawijaya, 1 Juni 2016.
3. Identifikasi berdasarkan gejala yang terlihat pada kucing.
4. Sistem ini akan mengidentifikasi 32 gejala dan 8 jenis penyakit yaitu Scabies, Gastritis, Helminthiasis, Rhinitis, Dermatophytosis, dermatitis, enteritis dan Otitis.

1.6 Sistematika pembahasan

Sistematika pembahasan ditunjukkan untuk memberikan gambaran dan uraian dari penyusunan tugas akhir secara garis besar yang meliputi beberapa bab, yaitu sebagai berikut:

BAB I : PENDAHULUAN

Bab 1 dijabarkan latar belakang penulis memilih topik skripsi ini, rumusan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan skripsi.

BAB II : LANDASAN KEPUSTAKAAN

Bab 2 membahas terkait kajian pustaka dari penelitian yang telah dilakukan sebelumnya dan dasar teori yang diperlukan untuk mendukung penelitian klasifikasi gejala penyakit kucing dengan menggunakan metode *K-Means Clustering* dan *Naive Bayes Classifier*.

BAB III : METODOLOGI

Bab 3 membahas tentang data penelitian, perancangan sistem, perhitungan manual, perancangan antarmuka dan rancangan pengujian sistem.

BAB IV : IMPLEMENTASI

Bab 4 membahas tentang spesifikasi perangkat lunak, Implementasi system dan rancangan pengujian perangkat lunak yang dibangun

BAB V : PENGUJIAN DAN ANALISIS

Bab 5 membahas tentang pengujian serta analisis terhadap aplikasi yang dibangun.

BAB VI : PENUTUP

Bab 6 berisi kesimpulan dari seluruh rangkaian penelitiandan saran untuk kemungkinan pengembangan lebih lanjut

BAB 2 LANDASAN KEPUSTAKAAN

Bab ini berisi teori - teori yang digunakan dalam proposal skripsi ini. Kajian pustaka membahas mengenai paper pendukung dan penelitian – penelitian yang sebelumnya telah ada. Sedangkan untuk dasar teori yang dibutuhkan ialah teori tentang klasifikasi, metode *k-means clustering* dan *naïve bayes classifier*, *K-Means Naïve Bayes* (KMNB) dan tentang penyakit pada kucing.

2.1 Kajian Pustaka

Dalam proposal skripsi ini, pada Kajian Pustaka akan dijabarkan mengenai penelitian-penelitian yang relevan terhadap judul yang diajukan. Penelitian pertama dengan judul *A K-Means and Naïve Bayes Learning Approach For Better Intrusion Detection* yang dilakukan oleh Z. Muda dkk. Dalam penelitian ini pendekatan K-Means dan Naïve Bayes Classifier (KMNB) dilakukan untuk melakukan deteksi intrusi. Hasil penelitian menunjukkan bahwa KMNB secara signifikan meningkatkan akurasi, tingkat deteksi, dan *false alarm* sampai 99.6, 99.8 dan 0,5% (Z. Muda dkk, 2011).

Penelitian terkait yang kedua adalah *K-Means Clustering and Naïve Bayes Classifier For Categorization Of Diabetes Patients*. Penelitian tersebut untuk kategorisasi penderita diabetes yang dilakukan oleh Pandeewari dkk. Dalam penelitian tersebut menyimpulkan integrasi *clustering* dan teknik klasifikasi membantu dalam mengidentifikasi data set yang besar. Teknik ini memberikan hasil yang lebih akurat daripada teknik klasifikasi sederhana (Pandeewari dkk, 2015).

Penelitian selanjutnya dengan judul *A Modified K-Means with Naïve Bayes (KMNB) Algorithm for Breast Cancer Classification* oleh Dian Eka Ratnawati dkk. Penelitian ini mengusulkan bahwa metode K-Means Naive Bayes (KMNB) dapat meningkatkan hasil klasifikasi pada data kanker payudara. Dengan melakukan modifikasi pada centroid awal dengan Fang. Keakuratan metode yang diusulkan adalah 95%, dan dapat mengurangi kesalahan sekitar 50%. Penelitian ini

menyatakan bahwa metode KMNB cukup menjanjikan dan mampu meningkatkan prediksi penyakit kanker payudara (Ratnawati dkk, 2018).

Penelitian mengenai metode *K-Means Clustering dan Naïve Bayes Classifier* dilakukan oleh Anggreyni adalah *Implementasi Metode Clustering Untuk Klasifikasi Kanker Payudara Menggunakan Algoritme Naïve Bayes dan K-Means*. Pada penelitian ini digunakan *clustering* K-Means dan Naïve Bayes. Pada K-Means dilakukan pengelompokan data. Kemudian hasil ‘mungkin’ dicari peluangnya dengan menggunakan Naïve Bayes. Penelitian ini membuktikan penggunaan K-Means dan Naïve Bayes pada dataset kanker payudara memberikan akurasi tertinggi sebesar 97,72% dan akurasi terendah 96,02% (Anggreyni, 2015).

Penelitian terkait yang keempat adalah *Penentuan Penerima Bantuan Ternak Menggunakan Algoritme K-Means & Naïve Bayes*, dilakukan oleh Mohamad Fadel Asikin. Penggabungan teknik klastering dan klasifikasi pada penelitian ini dianggap mampu memberikan hasil klasifikasi yang akurat pada penerimaan bantuan ternak. Pengelompokan data yang sama awal tahap dilakukan dengan pendekatan K-Means sebagai komponen pra-klasifikasi. Pendekatan selanjutnya menggunakan Naïve Bayes untuk klasifikasi data berdasarkan kategori diterima atau tidaknya bantuan ternak. Berdasarkan hasil pengujian dari penelitian ini, hasil pengelompokan menggunakan metode KMNB terbukti memberikan akurasi tertinggi sebesar 100% sedangkan K-Means konvensional memiliki akurasi sebesar 95.91 (Asikin, 2017).

2.2 Klasifikasi

Klasifikasi yaitu melakukan prediksi label suatu kelas untuk titik berlabel yang diberikan (J. Zaki & Meira.JR, p.29). Adapun pendapat lain mengenai *Classification* yaitu metode yang sering digunakan dalam konsep data mining untuk menyelesaikan masalah di lapangan. Klasifikasi melakukan pemrosesan untuk menemukan sebuah model atau fungsi yang menjelaskan dan mencirikan konsep atau kelas data, untuk kepentingan tertentu yaitu

menempatkan instans (object-object) baru (dengan label yang tak diketahui sebelumnya) ke dalam kelompok atau kelas masing-masing (Defiyanti Sofi, 2015).

2.3 K-Means Clustering dan Naïve Bayes Classifier

2.3.1 K-Means Clustering

K-Means merupakan metode pengelompokan yang paling terkenal dan banyak digunakan di berbagai bidang karena sederhana dan mudah diimplementasikan. K-Means merupakan metode *clustering* secara *partitioning* yang memisahkan data ke dalam kelompok yang berbeda. K-means merupakan salah satu metode pengelompokan data nonhirarki yang berusaha membagi data yang ada ke dalam bentuk dua atau lebih kelompok (Supranto, 2004). Metode ini dikembangkan oleh Mac Queen pada tahun 1967.

Tujuan dari pengelompokan data ini adalah untuk meminimalisasikan fungsi objektif dalam proses pengelompokan, yang pada umumnya berusaha meminimalisasikan ragam di dalam suatu kelompok dan memaksimalkan ragam antar kelompok. Dasar algoritme K-means dapat dilihat pada persamaan 2-1 :

1. Diberikan nilai k sebagai jumlah klaster yang ingin dibentuk.
2. Bangkitkan k *centroid* (titik pusat klaster) awal secara random.
3. Hitung jarak setiap data ke masing-masing pusat klaster yaitu menggunakan *Euclidean Distance* pada persamaan 2-1:

$$d(x,y)=|x,y|^2=\sqrt{\sum_{i=1}^n(x_i-y_i)^2} \quad (2-1)$$

Dimana:

$d(x,y)$ = ukuran ketidakmiripan

x = (x_1, x_2, \dots, x_j) adalah variabel data.

y = (y_1, y_2, \dots, y_j) adalah variabel pada titik pusat.

4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan pusatnya.

5. Tentukan posisi pusat klaster baru (C_{kj}) dengan cara menghitung nilai rata-rata dari data-data yang ada pada pusat klaster yang sama.

$$C_{kj} = \frac{x_{1j} + x_{2j} + \dots + x_{aj}}{a} \quad j = 1, 2, \dots, 7 \quad (2-2)$$

C_{kj} = pusat klaster ke-k pada variabel ke-j

a = banyak data pada klaster ke-k

2.3.2 Naïve Bayes Classifier

Naive Bayes Classifier (NBC) merupakan salah satu metode klasifikasi dimana berdasarkan konsep pada teorema Bayes. Metode ini memiliki ciri utama yaitu NBC mempunyai asumsi yang sangat kuat atau naïf. Sebelum menjelaskan NBC berikut konsep teorema Bayes yang mana sebagai dasar dari metode NBC. Pada teorema Bayes, apabila terdapat dua kejadian yang saling bebas (misalnya A dan B), maka perumusan dari teorema Bayesnya yaitu (Natalius, 2011):

$$P(A|B) = \frac{P(A)}{P(B)} P(B|A) \quad (2-3)$$

Penjelasan dari formula diatas adalah sebagai berikut :

$P(A|B)$: Probabilitas akhir bersyarat (*conditional probability*) suatu hipotesis A terjadi jika diberikan bukti (*evidence*) B terjadi.

$P(B|A)$: Probabilitas suatu bukti B terjadi akan mempengaruhi hipotesis A

$P(A)$: Probabilitas awal (*Priori*) hipotesis A terjadi tanpa memandang bukti apapun

$P(B)$: Probabilitas awal (*Priori*) bukti B terjadi tanpa memandang hipotesis/bukti yang lain.

Adapun pengembangan metode NBC mengingat berlakunya hukum probabilitas total, sehingga menjadi seperti berikut ini (Natalius, 2011):

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^n P(A_i|B)} \quad (2-4)$$

Penjelasan dari formula diatas sebagai berikut :

$P(A_j|B)$: Probabilitas akhir bersyarat (*conditional probability*) suatu hipotesis A_j terjadi jika diberikan bukti (*evidence*) B terjadi.

$P(B|A_j)$: Probabilitas suatu *evidence* B terjadi, jika hipotesis A_j terjadi.

$P(A_j)$: Probabilitas hipotesis A_j terjadi tanpa memandang *evidence* yang lain.

n : Jumlah hipotesis yang terjadi.

Untuk menjelaskan teorema *Naive Bayes*, hal yang perlu diketahui dalam sebuah proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas sebuah kelas yang tepat dari sampel yang dianalisis. Sehingga teorema Bayesnya sebagai berikut (Natalius, 2011):

$$P(C|F_1, \dots, F_n) = \frac{P(C)P(F_1, \dots, F_n|C)}{P(F_1, \dots, F_n)} \quad (2-5)$$

Penjelasan dari formula diatas sebagai berikut :

$P(C|F_1, \dots, F_n)$: Probabilitas *Posterior* bersyarat (*conditional probability*) suatu hipotesis C terjadi jika diberikan *evidence*/bukti F_1, \dots, F_n terjadi.

$P(F_1, \dots, F_n|C)$: Probabilitas awal (*Prior*) *evidence* suatu *evidence* F_1, \dots, F_n yang dipengaruhi hipotesis C .

$P(C)$: Probabilitas hipotesis C terjadi tanpa memandang *evidence* yang lain.

n : Jumlah hipotesis yang terjadi.

Dimana variabel C merepresentasikan kelas, sedangkan untuk variabel F_1, \dots, F_n merepresentasikan karakteristik-karakteristik yang digunakan dalam proses klasifikasi. Persamaan di atas menjelaskan peluang masuknya sampel dengan karakteristik-karakteristik tertentu dalam kelas C (*Posterior*) adalah

peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut *Prior* dimana perhitungan *Prior* sendiri yaitu probabilitas F_1, \dots, F_n dalam kelas C dibagi dengan jumlah probabilitas F_1, \dots, F_n dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas C (disebut dengan *Likelihood*), kemudian dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (*evidence*). Maka dari itu penulisan rumus sederhananya sebagai berikut (Natalius, 2011):

$$Posterior = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}} \quad (2-6)$$

Nilai *evidence* selalu bernilai tetap untuk setiap kelas pada sebuah sampel. Nilai *Posterior* tersebut nantinya akan dibandingkan dengan nilai-nilai *Posterior* kelas lainnya, hal ini untuk menentukan ke kelas apa sebuah sampel tersebut akan diklasifikasikan. Adapun proses perhitungan *Likelihood* jika datanya berupa kategori (diskret) ataupun nominal (*kontinue*), sebagai berikut (Meilani D B, 2014) :

a. Data Kategori

Data kategori adalah data yang sifatnya tetap dan tidak mengalami perubahan nilai. Berikut persamaan yang digunakan untuk menghitung data jenis kategori :

$$\frac{\text{jumlah data milik kelas C}}{\text{Total data sampel}} \quad (2-7)$$

Salah satu contoh atribut yang bersifat kategori adalah atribut jenis kelamin yaitu laki-laki dan perempuan.

b. Data Kontinyu

Data Kontinyu adalah jenis tipe data yang nilainya dapat berubah ubah. Biasanya data yang berjenis angka. Berikut rumus yang digunakan untuk menghitung probabilitas sebuah data yang bersifat kontinyu.

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (2-8)$$

$$\sigma = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2 \quad (2-9)$$

$$f(w) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(w-\mu)^2}{2\sigma^2}} \quad (2-10)$$

$f(w)$ merupakan nilai kemunculan probabilitas untuk data yang mempunyai nilai (w) . Salah satu contoh atribut yang bersifat kontinyu adalah atribut nilai ujian, dimana setiap siswa mempunyai nilai ujian yang berbeda-beda. Dana parameter μ adalah rata rata dan nilai π (phi) bernilai 3,14.

2.4 KMNB (K-Means Naïve Bayes)

K-Means Naïve Bayes (KMNB) dibentuk oleh kombinasi teknik *clustering* dan klasifikasi. Teknik *clustering* K-Means digunakan untuk mengelompokkan data berdasarkan kategori. Selanjutnya untuk data yang belum terkategori dalam pengelompokan akan diklasifikasikan ke dalam kategori yang ada menggunakan pengklasifikasian *Naïve Bayes*. Dalam hal ini data yang belum masuk klaster dengan algoritme *K-Means* akan diklasifikasikan menggunakan algoritme *Naïve Bayes*.

Prosedur yang digunakan untuk pengklasifikasian data dapat dilihat pada langkah-langkah berikut (Anggreyni, 2015):

1. Masukkan dataset.
2. Kelompokkan data dengan *K-Means* dan kembali dengan hasil kategori tertentu dan mungkin (belum masuk klaster kategori tertentu).
3. Jika hasil adalah dari kategori scabies maka keluaran adalah kategori scabies. jika hasil mungkin maka dilanjutkan ke langkah selanjutnya.
4. Data yang belum masuk klaster kategori tertentu (mungkin) akan diklasifikasikan dengan *Naïve Bayes*.

K-Means adalah salah satu metode yang paling sederhana dan populer dari algoritme pembelajaran yang bersifat arahan (*unsupervised*) untuk memecahkan masalah pengelompokan. Dalam *statistic* dan data mining, *K-Means* adalah sebuah metode analisis klaster yang bertujuan untuk partisi n

obyek ke dalam kelompok k didasari pada jarak yang disebut dengan apriori dan harus dihitung dengan langkah-langkah yang ada (Pandeewari dkk, 2015).

Kombinasi dari *clustering K-Means* dan pengklasifikasian *Naïve Bayes* menunjukkan peningkatan dibanding dengan pengklasifikasian tunggal dengan *Naïve Bayes*, karena ini dapat meningkatkan akurasi, nilai deteksi, dan mengurangi peringatan salah (Pandeewari dkk, 2015).

2.5 Penyakit Kucing

Kesehatan hewan peliharaan merupakan hal yang sangat penting, selain itu interaksi dengan manusia jika keadaan hewan kurang sehat maka dapat membahayakan kesehatan pemiliknya. Macam penyakit kucing yang biasa menjangkit kucing adalah sebagai berikut:

2.5.1 Scabies

Scabies atau kudis adalah salah satu penyakit kulit yang dapat menyebabkan kematian pada hewan kucing. Penyakit ini menimbulkan rasa gatal yang teramat sangat, lapisan kulit mengalami keratinasi, keropeng, ketombe, kutu, bulu rontok serta mengakibatkan tubuh kucing menjadi kurus. Jika penyakit scabies ini telah menginfeksi hewan kucing dalam stadium parah maka kucing akan mengalami penurunan sistem kekebalan tubuh secara drastis hingga dapat menyebabkan kematian. Menurut Drh. Naumi D.R.P selain menular kepada kucing lain, penyakit kulit scabies juga dapat menular kepada manusia. Walaupun tidak menyebabkan kematian pada manusia, penyakit scabies dapat menimbulkan rasa gatal yang cukup mengganggu, maka dari itu jenis penyakit ini harus dilakukan pengobatan dengan benar, cepat dan tepat oleh pemiliknya secara dini (Palguna David dkk, 2014).

2.5.2 Gastritis

Penyakit gastritis disebabkan lingkungan hewan peliharaan yang tidak bersih sehingga kemungkinan besar hewan memakan benda asing atau makanan yang mengandung bahan berbahaya sehingga menyebabkan anoreksia. Gastritis

menyebabkan iritasi kronis mukosa gastrik terjadi akibat respon inflamasi pada mukosa yang meluas hingga lapisan submukosa. Gastritis kronis sekunder terjadi akibat respon imun atau alergi yang berhubungan dengan stimulasi antigenik kronis. Salah satu gejala klinis yang diderita kucing jika mengindap penyakit ini yaitu mengalami vomitus (muntah bercampur empedu) dan berisi makanan yang belum tercerna dengan baik, abdomen keras, diare hingga tubuh kucing menjadi kurus (Triakoso N, 2006).

2.5.3 Helminthiasis

Helminthiasis merupakan penyakit yang disebabkan oleh cacing gastrointestinal. Cacing ini mengakibatkan infeksius pada inangnya, menghambat pertumbuhan. Gejala klinis yang disebabkan akan cacing ini yaitu terjadinya diare, perut buncit, muntah, hilang nafsu makan, dan terdapat cacing pada feses kucing. Helminthiasis hidup dalam saluran usus, dimana enzim pada pencernaan dapat melarutkan kulit telur dan melepaskan cacing. Kemudian cacing berkembangbiak sehingga menghasilkan telur lebih banyak lagi lalu menetas hingga melanjutkan siklus hidupnya. Meskipun enzim pencernaan dapat melarutkan kulit telur, namun enzim tersebut tidak akan membutuh ataupun berbahaya bagi cacing dewasa hal ini disebabkan cacing dewasa memiliki lapisan keratin diluar tubuh cacing. Cacing ini mempunyai struktur tubuh multiseluler dengan sistem organ. Terdapat kemungkinan besar pemberian obat-obatan tertentu dapat membunuh cacing dan mencegah reproduksi (Gustiar Rendra, 2011).

2.5.4 Rhinitis

Rhinitis merupakan salah satu penyakit yang menyerang saluran pernafasan, dimana terjadi infeksi yang melibatkan organ pernafasan seperti hidung, sinus, faring, trachea, bronchus, bronkiolus hingga ke paru-paru. Penyebab penyakit ini yaitu dapat berupa virus, bakteri, dan jamur. Penyakit rhinitis yang disebabkan oleh virus dapat menularkan secara langsung dari kucing ke kucing lain melalui kontak langsung dengan cairan terinfeksi dari mata,

hidung, mulut, makanan yang telah terinfeksi bahkan lewat udara. Virus ini dapat hidup dengan stabil selama 24 jam sampai 10 hari tergantung kondisi lingkungannya. Gejala klinis yang terjadi jika mengindap penyakit ini yaitu mulai dari pilek, bersin-bersin, hidung tersumbat, badan lemas, mata dan hidung berair, serta anoreksis (Putra Alimansyah, 2012).

2.5.5 Dermatitis

Dermatitis merupakan penyakit kulit berupa peradangan pada kulit sehingga menyebabkan bulu rontok, keropeng pada kulit, kulit menjadi kemerahan bahkan hingga lecet, jamur, dan munculnya lesi berminyak. Hal ini biasanya disebabkan alergi oleh berbagai macam, mulai dari lingkungan ataupun pakan. Gejala klinisnya selain menyebabkan gatal hewan yang terkena dermatitis akan selalu menggaruk-garuk dan menjilati seluruh bagian tubuhnya. Penyakit ini bersifat sementara dan kronik sehingga menyebabkan gangguan pada organ yang lain (Anonim, 2011). Salah satu penyebab penyakit ini adalah lingkungan seperti debu, serbuk bunga, spora, dan alergen lain-lain. Ada tiga hal yang harus menjadi pertimbangan untuk mendiagnosis atopik dermatitis yaitu (Triakoso N, 2013):

- ✓ Sejarah atau anamnesis
- ✓ Pemeriksaan fisik, ditemukan gejala-gejala yang mengarah
- ✓ Diferensial diagnosis, sudah telah dilakukan "*rule out*" terhadap penyakit-penyakit lain yang mirip dengan *atopic dermatitis*.
- ✓ Selain itu diperlukan untuk melakukan *skin test* bukan serum *allergy test* untuk mengetahui sumber alergennya.

2.5.6 Dermaphytosis

Dermaphytosis merupakan penyakit kulit yang disebabkan oleh jamur, penyakit ini pada umumnya bersifat zoonotik dan sangat tinggi penularannya. Pengobatan penyakit ini membutuhkan waktu lama disebabkan rata rata pasien mengalami infeksi sehingga membutuhkan biaya yang besar dan waktu pengobatan yang intens. Penyakit ini sering dianggap sebagai hasil diagnosis

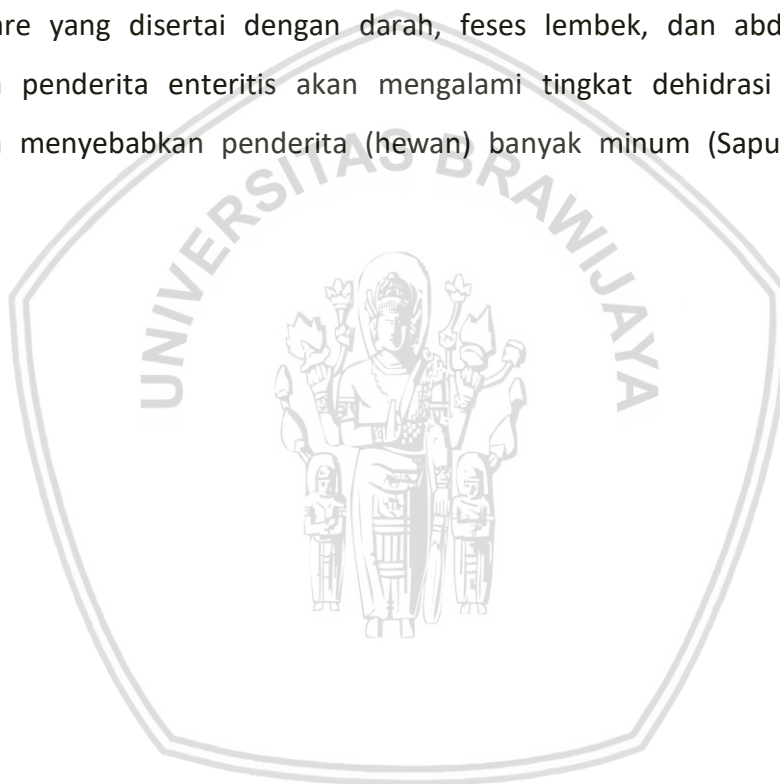
penyakit kulit biasa oleh para dokter hewan. Infeksi pada penyakit ini ditandai dengan kulit mengalami peradangan dan biasanya berbentuk seperti cincin sehingga disebut dengan istilah *ringworm dermatophyte*, mengalami gatal-gatal, bulu rontok, keropeng pada kulit, hingga kulit menjadi merah sampai lecet. Bulu kucing yang tebal menjadi salah satu habitat yang disukai oleh jamur penyebab penyakit ini, selain itu kondisi lingkungan dengan kelembaban yang tinggi juga merupakan faktor pendukungnya (Adzima Vhodzan, 2013).

2.5.7 Otitis

Penyakit otitis merupakan kondisi dimana terjadinya peradangan pada telinga kucing, hal ini disebabkan oleh bakteri, jamur, alergi, serta gangguan sistem kekebalan tubuh. Otitis dapat terjadi pada bagian telinga yaitu bagian luar, tengah, dan dalam. Jika kondisi penyakit terjadi pada telinga bagian dalam maka dapat menyebabkan hilangnya kemampuan mendengar kucing. Pengobatan akan penyakit otitis membutuhkan penanganan yang tepat, jika tidak maka bisa menyebabkan tumbuhnya polip dimana dapat tumbuh menjadi kanker dan menutupi saluran telinga. Apabila kucing telah terjangkit penyakit ini biasanya kucing akan merasa tidak nyaman serta sering kali mengoyang/menggelengkan kepalanya, mencakar-cakar telinganya hingga timbul gurat parah pada telinga, ataupun menggosok-gosokkan ke benda lain seperti dinding. Selain itu telinga kucing terdapat cairan hitam serta terdapat lilin dengan bau yang menyengat. Kucing mengalami anoreksia dan timbul keropeng pada kulit. Cakaran atau goyangan kepala yang terus menerus dalam jangka waktu lama dapat menyebabkan hematoma pada telinga (aural hematoma). Hematoma adalah penggumpalan atau penumpukan darah di telinga akibat pecahnya pembuluh darah yang terdapat pada daun telinga. Telinga yang mengalami hematoma terlihat dari tanda-tanda seperti bengkak, dan terasa hangat bila diraba dan terasa ada penumpukan cairan di bawah kulit telinga (Neno, 2016).

2.5.8 Enteritis

Penyakit enteritis merupakan suatu penyakit kucing yang disebabkan oleh virus *feline parvo*, dimana virus ini sangat cepat menular. Virus FPL ini menyerang segala umur kucing dan dapat menimbulkan kematian pada anak kucing hingga prosentase 75%. Virus FPL masuk dalam famili *Parvoviridae* yang menyerang jaringan pembentuk darah dan limfe, dan juga mukosa organ gastro intestinal sehingga menyebabkan penurunan jumlah leukosit dan enteritis. Adapun gejala klinis yang terlihat yaitu anoreksia selama periode inkubasi 2-7 hari, diare yang disertai dengan darah, feses lembek, dan abdomen sakit. Biasanya penderita enteritis akan mengalami tingkat dehidrasi yang tinggi sehingga menyebabkan penderita (hewan) banyak minum (Saputra Thomas, 2015).



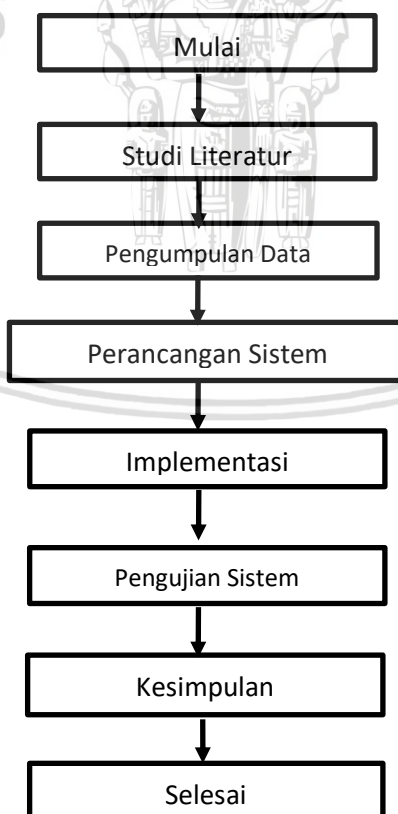
BAB 3 METODOLOGI

3.1 Jenis Penelitian

Penelitian yang dilakukan merupakan penelitian implementatif pembangunan. Penelitian jenis ini merupakan sebuah kegiatan penelitian dalam rangka membuat sebuah produk perangkat lunak dengan sistematis. Penelitian dimulai dari studi literatur, pengumpulan data, analisis kebutuhan, perancangan perangkat lunak berdasarkan analisis kebutuhan, konstruksi atau pembuatan perangkat lunak, serta pengujian hasil jadi produk perangkat lunak.

3.2 Metodologi Penelitian

Langkah-langkah yang akan dilakukan dalam penelitian ini dapat di lihat pada Gambar 3.1.



Gambar 3.1 Diagram Alir Metodologi Penelitian

3.2.1 Studi Literatur

Studi literatur mempelajari mengenai dasar teori yang digunakan untuk menunjang penulisan pengerjaan skripsi. Teori-teori pendukung penulisan serta pemahaman tentang skripsi diperoleh dari buku, jurnal, *e-book* dan penelitian sebelumnya yang berkaitan tentang topik skripsi ini. Studi literatur ini meliputi:

1. Kajian Pustaka
2. Klasifikasi
3. *K-Means Clustering*
4. *Naive Bayes Classifier*
5. *K-Means Naive Bayes (KMNB)*
6. Penyakit Kucing

3.2.2 Analisis Kebutuhan

Analisa Kebutuhan Sistem merupakan tahap menganalisis hal-hal yang dibutuhkan untuk menjalankan sistem, sehingga sistem dapat berjalan secara optimal.

3.2.3 Pengumpulan Data

Lokasi penelitian skripsi ini pada klinik hewan Universitas Brawijaya Malang. Variabel penelitian skripsi ini adalah penyakit yang menyerang kucing disertai gejala yang tampak berdasarkan hasil nilai perhitungan menggunakan metode *K-Means* dan *Naive Bayes Classifier*. Data berasal dari skripsi sebelumnya yang didapat dari hasil konsultasi dengan dokter hewan klinik hewan Universitas Brawijaya. Data diolah sedemikian rupa berdasarkan proses perhitungan metode *K-Means* dan *Naive Bayes*.

3.2.4 Perancangan Sistem

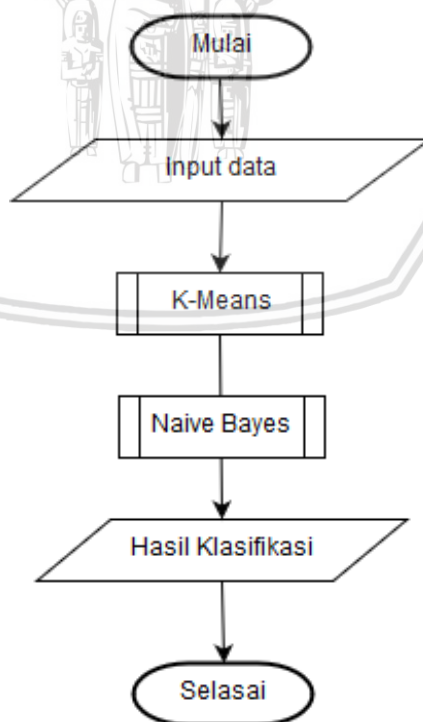
Perancangan sistem digunakan untuk merancang langkah kerja dari sistem yang akan dibangun secara menyeluruh. Pada pemecahan masalah untuk diagnosa penyakit pada kucing diperlukan tahap perancangan sistem dimana

perancangan sistem ini terdiri atas deskripsi umum sistem, perancangan model *K-Means* dan *Naïve Bayes*, perancangan *K-Means* dan *Naïve Bayes* dengan diagram alir, contoh perhitungan manual, perancangan antarmuka dan perancangan pengujian sistem.

3.2.4.1 Deskripsi Umum Sistem

Secara umum sistem yang dibuat akan menampilkan hasil keluaran dari metode *K-Means Clustering* dan *Naive Bayes Classifier* berupa penyakit. Proses klasifikasi yang dilakukan yaitu User melakukan input dataset dengan cara meng-*import* data dari tempat penyimpanan data. Proses selanjutnya adalah dilakukan proses perhitungan *K-Means Clustering* kemudian *Naive Bayes Classifier*. Adapun parameter kelas yang digunakan dalam perhitungan ini, menggunakan 8 kelas penyakit kucing yaitu scabies, gastritis, helminthiasis, rhinitis, dermaphytosis, dermatitis, enteritis, otitis, dan kelas mungkin serta menggunakan 32 gejala penyakit.

Secara keseluruhan alur proses sistem dapat dilihat pada Gambar 3.2.



Gambar 3.2 Alur Proses Sistem

3.2.4.2 Perancangan Sistem

Pada sub bab ini akan dibahas mengenai perancangan sistem yaitu perancangan sistem klasifikasi dengan melibatkan proses pengolahan data dimana dalam hal ini menggunakan metode *K-Means Clustering* dan *Naive Bayes Classifier*. Pada proses pengolahan data peneliti membagi proses pada sistem menjadi 2 bagian, yaitu proses *clustering* data dengan *K-Means* dan proses klasifikasi data dengan *Naive Bayes*.

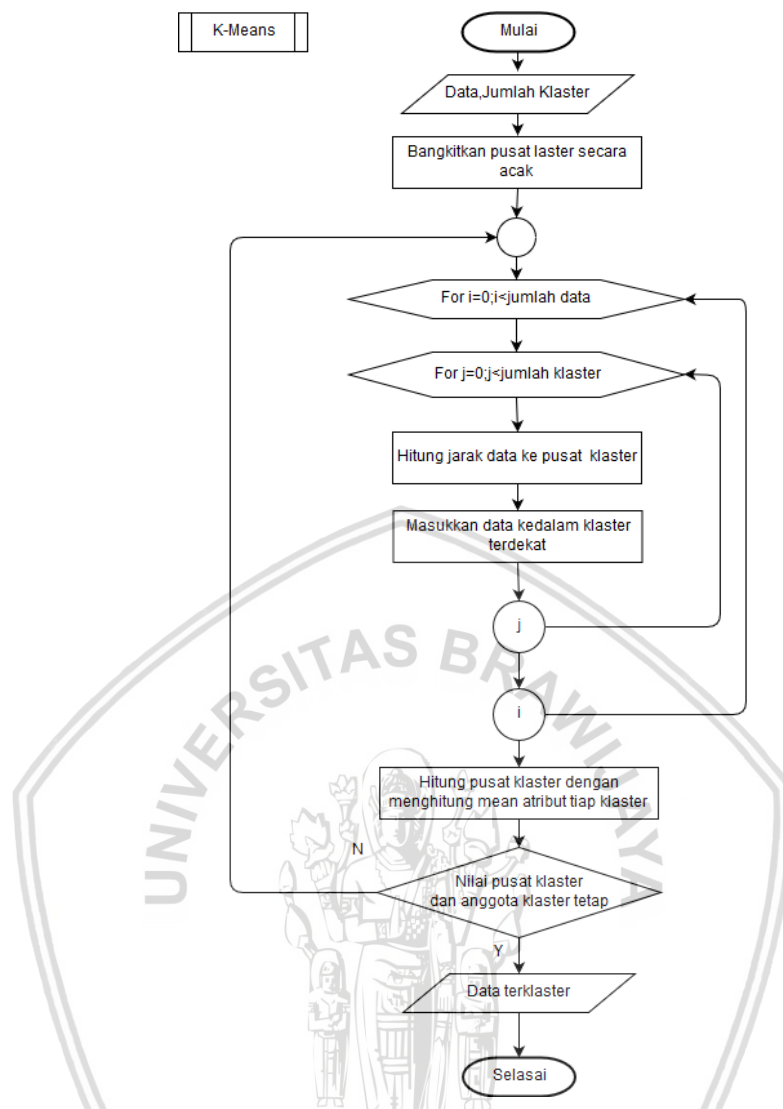
(a) *Clustering* Data dengan K-Means

Prinsip dari proses pengelompokan dengan menggunakan K-Means pada penelitian ini adalah membentuk sebanyak 9 kluster.

Proses ini dijabarkan sebagai berikut:

1. Masukan berupa data gejala penyakit kucing.
2. Menentukan nilai k .
3. Membangkitkan pusat kluster secara acak.
4. Melakukan perhitungan untuk jarak setiap data pada masing-masing pusat kluster.
5. Melakukan *update* kluster, yaitu memasukkan data ke dalam kelompok dengan jarak minimum ke pusat kluster berdasarkan jarak setiap data pada masing-masing pusat kluster.
6. Melakukan *update* pusat kluster, jika sudah periksa anggota kluster apakah berubah. Jika berubah akan kembali ke langkah 4, jika tidak ada perubahan maka proses *clustering* selesai.

Alur sistem pengelompokan K-Means dapat dilihat pada Gambar 3.3.



Gambar 3.3 Alur Sistem Pengelompokan K-Means

Perhitungan pada *clustering* data gejala penyakit kucing dengan K-Means yang pertama yaitu menentukan jumlah klaster. Pada penelitian ini menggunakan 9 klaster untuk menentukan jenis penyakit pada kucing, yaitu klaster 1 adalah scabies, klaster 2 adalah gastritis, klaster 3 adalah helminthiasis, klaster 4 adalah rhinitis, klaster 5 adalah dermatophytosis, klaster 6 adalah dermatitis, klaster 7 adalah enteritis, klaster 8 adalah otitis, dan klaster 9 adalah mungkin (tidak teridentifikasi penyakit oleh *K-Means*).

Proses pada *K-Means* diawali dengan masukan data gejala penyakit kucing, dan parameter *K-Means* yaitu jumlah klaster. Proses kedua, dengan

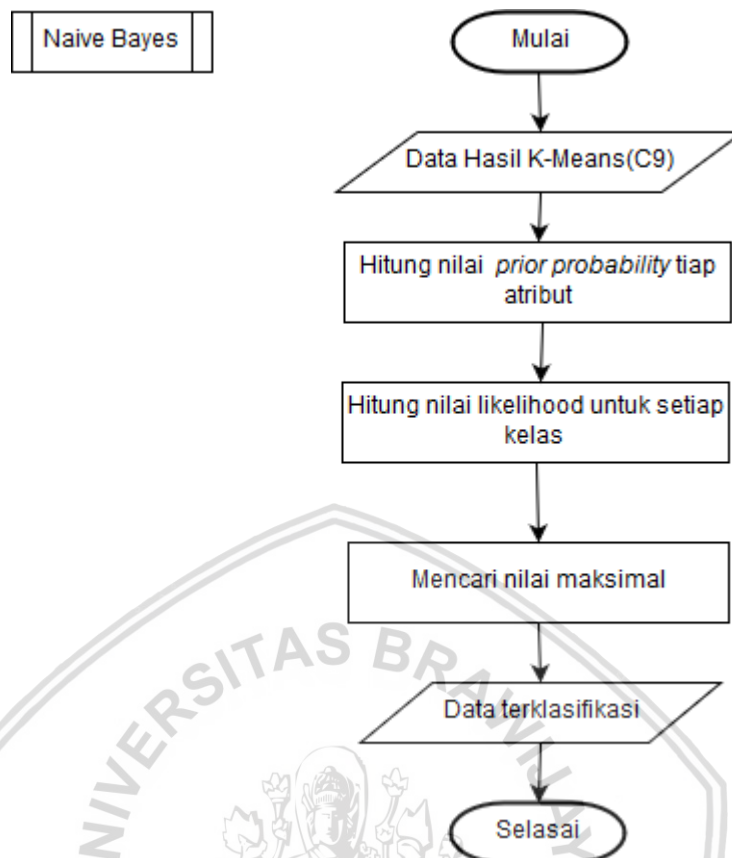
membangkitkan pusat klaster secara acak untuk dijadikan sebagai pusat klaster. Kemudian untuk setiap data dihitung jarak *euclidian* data ke pusat klaster. Data akan dimasukkan ke dalam klaster dengan memilih nilai jarak *euclidian* yang terdekat (terkecil). Selanjutnya jika seluruh data masuk ke dalam klaster, maka akan dihitung pusat klaster baru dengan cara menghitung rata-rata (mean) dari seluruh data dalam klaster tersebut dan menghitung jarak *euclidian*-nya. Selanjutnya akan dilihat apakah pusat klaster dan anggota klaster tetap. Jika mengalami perubahan maka dilakukan kembali perhitungan jarak data ke pusat klaster. Jika sudah tidak mengalami perubahan (*konvergen*), maka proses iterasi selesai dan didapatkan keanggotaan masing-masing data dengan klasternya.

(b) Klasifikasi Data dengan Naïve Bayes

Pada proses ini sistem akan melakukan klasifikasi menggunakan metode Naive Bayes. Proses tahapannya adalah sebagai berikut :

1. Masukan data latih.
2. Melakukan perhitungan nilai *prior probability* dari setiap kelas untuk satu atribut data gejala penyakit kucing.
3. Menghitung probabilitas setiap atribut untuk semua kelas.
4. Melakukan perhitungan nilai likelihood untuk setiap kelas.
5. Mencari nilai maksimal untuk mencari kelas data.

Alur proses Naïve Bayes dapat dilihat pada Gambar 3.4.



Gambar 3.4 Alur Klasifikasi Naïve Bayes

Proses pertama yaitu menginputkan data hasil K-Means (C9), setelah itu menghitung nilai *prior probability* dari setiap atribut. Setelah nilai didapatkan, di hitung nilai *likelihood* untuk setiap kelas. Tahap berikutnya adalah mencari nilai probabilitas yang paling maksimum untuk setiap kelas yang dihasilkan.

3.2.4.3 Perhitungan Manual

(a) K-MEANS

Pada perhitungan K-Means diperlukan data latih yang diambil dari dataset gejala kucing sejumlah 40 data dengan 32 gejala yang dibutuhkan untuk proses pengklasifikasian. Contoh dari data latih yang digunakan dalam proses K-Means seperti pada Tabel 3.1.

Tabel 3.1 Contoh Data Gejala Kucing

Data ke	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29	G30	G31	G32	Diagnosa	
1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	scabies	
2	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	scabies	
3	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	scabies	
4	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	scabies
5	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	scabies
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	rhinitis
7	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	rhinitis
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	rhinitis
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	rhinitis
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	rhinitis
11	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	otitis
12	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	otitis
13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	otitis
14	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	otitis
15	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	otitis
16	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	helminthiasis
17	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	helminthiasis
18	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	helminthiasis
19	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	helminthiasis
20	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	helminthiasis
21	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	gastritis
22	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	gastritis

Data ke	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29	G30	G31	G32	Diagnosa
23	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	gastritis
24	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	gastritis
25	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	gastritis
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	enteritis
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	enteritis
28	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	enteritis
29	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	enteritis
30	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	enteritis
31	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	dermatophytosis
32	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	dermatophytosis
33	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	dermatophytosis
34	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	dermatophytosis
35	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	dermatophytosis
36	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	dermatitis
37	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	dermatitis
38	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	dermatitis
39	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	dermatitis
40	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	dermatitis

Tabel 3.2 Kode Gejala

Gelaja	Kode Gejala	Gejala	Kode Gejala
Keratinisasi	G1	Hidung Tersumbat	G17
Gatal gatal	G2	Badan lemas	G18
Keropeng	G3	Mata berair	G19
Ketombe	G4	Hidung berair	G20
Kutuan	G5	Ringworm pada kulit	G21
Kurus	G6	Kulit kemerahan sampai lecet	G22
Bulu Rontok	G7	Jamuran	G23
Anoreksia	G8	lesi berminyak pengganti	G24
Abdomen keras	G9	Guratan parah pada telinga	G25
Muntah	G10	Adanya cairan hitam keluar telinga	G26
Diare	G11	Telinga terdapat lilin dan bau	G27
Perut buncit	G12	Diare campur darah	G28
Hilang nafsu makan	G13	Feses lembek	G29
Ada cacing di feses	G14	Minum banyak	G30
Pilek	G15	Abdomen sakit	G31
Bersin bersin	G16	Vaksinasi	G32
Nilai 0 = Tidak bergejala			
Nilai 1 = Memiliki gejala			

Pada Tabel 3.1 terdapat data gejala penyakit kucing sejumlah 40 data ($n=40$). Sedangkan atribut yang dimiliki adalah 32 atribut ($m=32$). Dari 40 data tersebut akan dikelompokkan menjadi 9 *cluster* ($c=9$). Langkah pertama yaitu dilakukan proses pembangkitan pusat *cluster* (*centroid*) awal yang digunakan dalam proses *clustering* di setiap atributnya. Pemilihan pusat *cluster* dilakukan secara acak, misalkan terpilih data yang ditunjukkan pada Tabel 3.3 sebagai pusat klaster awal.

Tabel 3.3 Pusat kluster awal

Pusat Cluster Awal																																	
	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29	G30	G31	G32	
C1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
C2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
C4	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C5	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
C7	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C8	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
C9	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1

Selanjutnya dilakukan perhitungan jarak masing-masing data dengan pusat kluster dengan menggunakan rumus jarak *eucliden* (persamaan 2-1) dengan keterangan x adalah data, c adalah kluster.

$$d(x1,c1) = \sqrt{(0-1)^2 + (1-0)^2 + (1-0)^2 + (0-1)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-1)^2}$$

$$= 1.41$$

$$d(x1,c2) = \sqrt{(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2}$$

$$= 1.41$$

$$d(x1,c3) = \sqrt{(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2}$$

$$= 1.41$$

$$d(x1,c4) = \sqrt{(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2}$$

$$= 1.41$$

$$d(x1,c5) = \sqrt{(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + (0-1)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2}$$

$$= 1.41$$

$$d(x1,c6) = \sqrt{(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-1)^2 + (0-1)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2}$$

$$= 1.41$$

$$d(x1,c7) = \sqrt{(0-0)^2 + (1-1)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2}$$

$$= 1.00$$

$$d(x1,c8) = \sqrt{(0-0)^2 + (1-0)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2}$$

$$= 1.00$$

$$d(x1,c9) = \sqrt{(0-0)^2 + (1-0)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-1)^2}$$

$$= 1.00$$

...

$$d(x40,c1) = \sqrt{(0-1)^2 + (0-0)^2 + (1-0)^2 + (0-1)^2 + (0-1)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2}$$

$$= 1.41$$

$$d(x40,c2) = \sqrt{(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2}$$

$$= 2.00$$

$$d(x40,c3) = \sqrt{(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2}$$

$$= 2.00$$

$$d(x_{40},c_4) = \sqrt{(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2}$$

$$= 2.00$$

$$d(x_{40},c_5) = \sqrt{(0-0)^2 + (1-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2}$$

$$= 2.00$$

$$d(x_{40},c_6) = \sqrt{(0-0)^2 + (1-1)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-1)^2 + (0-1)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2}$$

$$= 2.00$$

$$d(x_{40},c_7) = \sqrt{(0-0)^2 + (1-1)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2}$$

$$= 1.73$$

$$d(x_{40},c_8) = \sqrt{(0-0)^2 + (1-0)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2}$$

$$= 1.73$$

$$d(x_{40},c_9) = \sqrt{(0-0)^2 + (1-0)^2 + (1-1)^2 + (0-0)^2 + (0-0)^2 + \dots + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2}$$

$$= 0.00$$

Hasil keanggotaan tiap kluster dari perhitungan jarak *eucledian* dari iterasi ke=1 menggunakan 9 pusat kluster dapat dilihat pada Tabel 3.4.

Tabel 3.4 Hasil Perhitungan Jarak *Euclidian* Iterasi ke-1

Data ke-	C1	C2	C3	C4	C5	C6	C7	C8	C9	Min	Kluster
1	1.41	1.41	1.41	1.41	1.41	1.41	1.00	1.00	1.00	1.00	C7
2	0.00	2.24	2.24	2.24	2.24	2.24	2.24	2.24	1.73	0.00	C1
3	1.73	1.73	1.41	1.41	1.73	1.73	1.41	1.41	1.00	1.00	C9
4	2.00	2.00	2.00	2.00	1.73	2.00	1.41	1.73	1.41	1.41	C7
5	1.00	1.73	1.73	1.73	1.73	1.73	1.73	1.41	1.41	1.00	C1
6	2.00	0.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	0.00	C2
7	1.73	1.00	1.41	1.41	1.73	1.73	1.73	1.73	1.73	1.00	C2
8	1.41	0.00	1.41	1.41	1.41	1.41	1.41	1.41	1.41	0.00	C2
9	1.41	0.00	1.41	1.41	1.41	1.41	1.41	1.41	1.41	0.00	C2
10	1.41	0.00	1.41	1.41	1.41	1.41	1.41	1.41	1.41	0.00	C2
11	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.00	1.00	1.00	C8
12	1.73	1.73	1.73	1.73	1.73	1.73	1.41	1.41	1.00	1.00	C9
13	1.41	1.41	0.00	1.00	1.41	1.41	1.41	1.41	1.41	0.00	C3
14	1.73	1.73	1.73	1.73	1.73	1.73	1.73	1.41	1.41	1.41	C8
15	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	C1
16	1.73	1.73	1.73	1.00	1.00	1.73	1.73	1.73	1.73	1.00	C4
17	1.73	1.73	1.41	0.00	1.00	1.73	1.73	1.73	1.73	0.00	C4
18	1.41	1.41	1.00	1.00	1.41	1.41	1.41	1.41	1.41	1.00	C3
19	1.41	1.73	1.73	1.00	1.00	1.73	1.73	1.73	1.41	1.00	C4
20	1.00	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.00	1.00	C1
21	1.73	1.73	1.73	1.00	1.00	1.73	1.73	1.73	1.73	1.00	C4
22	1.73	1.73	1.73	1.00	0.00	1.73	1.73	1.73	1.73	0.00	C5
23	1.41	1.41	1.41	1.00	1.00	1.41	1.41	1.41	1.41	1.00	C4
24	1.73	1.73	1.73	1.41	1.00	1.73	1.73	1.73	1.73	1.00	C5
25	1.73	2.00	1.73	1.41	1.41	2.00	2.00	2.00	1.73	1.41	C4
26	1.41	1.41	1.41	1.41	1.41	0.00	1.41	1.41	1.41	0.00	C6
27	1.73	1.73	1.73	1.73	1.73	1.41	1.73	1.73	1.73	1.41	C6
28	2.00	2.00	1.73	1.41	1.73	1.73	2.00	2.00	2.00	1.41	C4
29	1.73	1.73	1.73	1.41	1.41	1.00	1.73	1.73	1.73	1.00	C6
30	1.73	2.00	1.73	1.73	2.00	1.73	2.00	2.00	1.73	1.73	C1
31	1.73	1.73	1.73	1.73	1.73	1.73	1.41	1.41	1.00	1.00	C9
32	1.41	2.00	2.00	2.00	2.00	2.00	1.73	2.00	1.73	1.41	C1
33	1.41	1.73	1.73	1.73	1.73	1.73	1.00	1.73	1.41	1.00	C7
34	1.41	1.41	1.41	1.41	1.41	1.41	0.00	1.41	1.00	0.00	C7
35	1.41	1.73	1.73	1.73	1.73	1.73	1.41	1.41	1.00	1.00	C9
36	1.41	2.00	2.00	2.00	2.00	2.00	1.73	1.73	1.00	1.00	C9
37	1.73	1.73	1.73	1.73	1.73	1.73	1.73	0.00	1.41	0.00	C8
38	1.41	1.73	1.73	1.73	1.73	1.73	1.73	1.00	1.00	1.00	C8
39	1.41	1.73	1.73	1.73	1.73	1.73	1.41	1.41	1.00	1.00	C9
40	1.41	2.00	2.00	2.00	2.00	2.00	1.73	1.73	0.00	0.00	C9

Dari Tabel 3.4 dapat diketahui bahwa anggota masing-masing kluster adalah sebagai berikut:

C1 = data ke-2, data ke-5, data ke-15, data ke-20, data ke-32

C2 = data ke-6, data ke-7, data ke-8, data ke-9, data ke-10,

C3 = data ke-13, data ke-18

C4 = data ke-16, data ke-17, data ke-19, data ke-21, data ke-23, data ke-25, data ke-28

C5 = data ke-22, data ke-24

C6 = data ke-26, data ke-27, data ke-29

C7 = data ke-1, data ke-4, data ke-33, data ke-34

C8 = data ke-11, data ke-14, data ke-37, data ke-38

C9 = data ke-3, data ke-12, data ke-35, data ke-36, data ke-39, data ke-40

Penentuan untuk keanggotaan kluster diketahui dari membandingkan hasil perhitungan jarak disetiap kluster dan diambil dengan jarak terkecil, misalnya sebagai berikut:

Jarak x ke C1 = 0,00 (terkecil)

Jarak x ke C2 = 2,24

Jarak x ke C3 = 2,24

Jarak x ke C4 = 2,24

Jarak x ke C5 = 2,24

Jarak x ke C6 = 2,24

Jarak x ke C7 = 2,24

Jarak x ke C8 = 2,24

Jarak x ke C9 = 1,73

Dari data diatas, maka x masuk dalam keanggotaan kluster 1 (C1).

Setelah mendapatkan seluruh anggota tiap kluster, selanjutnya dilakukan perhitungan untuk pusat kluster baru dari iterasi ke-1. Dengan cara menghitung mean data anggota kluster dari setiap kluster. Hasil perhitungan pusat kluster baru ditunjukkan pada Tabel 3.5 berikut ini:

Tabel 3.5 Perhitungan Kluster Iterasi ke - 1

Pusat Cluster Awal	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29	G30	G31	G32
C1	0.2	0.2	0.2	0.4	0.6	0	0	0	0	0	0	0	0.2	0.2	0	0	0	0	0	0	0.2	0	0.4	0	0	0	0	0.2	0	0	0.2	0.6
C2	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0	0.6	0.8	0.6	0	0.4	0	0	0	0	0	0	0	0	0	0	0	0	0
C3	0	0	0	0	0	0	0	0	0	0	0	0.5	1	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0	0
C4	0	0	0	0	0	0.2	0	0.2	0	1	0.5	0	0.5	0.2	0	0	0	0	0	0	0	0	0.2	0	0	0	0	0.2	0	0	0.2	0.2
C5	0	0	0	0	0	0.7	0	0	0.7	1	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C6	0	0	0	0	0	0	0	0	0	0.3	0	0	0	0	0	0	0	0.3	0	0	0	0	0	0	0	0	0	1	0.7	0	0.3	0
C7	0	1	0.5	0.3	0	0.3	0.8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C8	0	0	0.8	0	0	0	0	0.6	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.2	0.2	0.2	0	0.2	0	0	0	0	0
C9	0	0	0.7	0	0.1	0	1	0.1	0	0	0	0	0.1	0	0	0	0	0	0	0	0	0	0.4	0.3	0	0	0.1	0	0	0	0	0.3

Selanjutnya dilakukan perhitungan iterasi ke-2 karena anggota kluster masih mengalami perubahan. Iterasi dilakukan sampai dengan iterasi ke-3 dengan cara yang sama. Iterasi berhenti jika sudah tidak terjadi perubahan pada data.

Hasil iterasi ke-2 untuk membuktikan bahwa iterasi ke-2 dan ke-3 memiliki hasil sama sehingga iterasi berhenti, ditunjukkan pada Tabel 3.6.

Tabel 3.6 Hasil Perhitungan Jarak *Eucliden* Iterasi ke - 2

Data ke-	C1	C2	C3	C4	C5	C6	C7	C8	C9	Min	Kluster
1	1.1	1.4	1.4	1.4	1.4	1.4	0.5	1	1	0.5	C7
2	1.3	2.2	2.2	2.1	2.2	2.2	2.1	2.2	1.9	1.3	C1
3	1.5	1.6	1.4	1.5	1.7	1.7	1.1	1.4	0.9	0.9	C9
4	1.8	2	2	1.9	1.8	2	0.9	1.7	1.4	0.9	C7
5	1.1	1.7	1.7	1.7	1.7	1.7	1.3	1.4	1.3	1.1	C1
6	2	0.8	2	2	2	2	2	2	2	0.8	C2
7	1.6	0.9	1.4	1.5	1.7	1.7	1.7	1.7	1.7	0.9	C2
8	1.4	0.7	1.4	1.4	1.4	1.4	1.4	1.4	1.4	0.7	C2
9	1.4	0.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	0.4	C2
10	1.4	0.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	0.4	C2
11	1.3	1.4	1.4	1.3	1.4	1.4	1.1	0.4	0.9	0.4	C8
12	1.6	1.7	1.7	1.7	1.7	1.7	1.1	1.3	0.9	0.9	C9
13	1.3	1.3	0.5	1.1	1.4	1.4	1.4	1.4	1.3	0.5	C3
14	1.6	1.7	1.7	1.6	1.7	1.7	1.5	0.9	1.3	0.9	C8
15	1.4	1.4	1.4	1.3	1.4	1.4	1.4	0.9	1.2	0.9	C8
16	1.6	1.7	1.7	1	1.1	1.6	1.7	1.7	1.7	1	C4
17	1.6	1.6	1.4	0.7	1.1	1.6	1.7	1.7	1.7	0.7	C4
18	1.3	1.3	0.5	1.1	1.4	1.4	1.4	1.4	1.3	0.5	C3
19	1.5	1.7	1.7	1	1.1	1.6	1.7	1.6	1.5	1	C4
20	1	1.4	1.4	1.2	1.4	1.4	1.4	1.3	1.2	1	C1
21	1.7	1.7	1.7	1.1	0.5	1.6	1.7	1.7	1.7	0.5	C5
22	1.7	1.7	1.7	1	0.5	1.6	1.6	1.7	1.7	0.5	C5
23	1.4	1.4	1.4	0.8	1	1.2	1.4	1.1	1.3	0.8	C4
24	1.7	1.7	1.7	1.3	0.5	1.6	1.6	1.7	1.7	0.5	C5
25	1.7	1.9	1.7	1.3	1.5	1.9	1.9	2	1.8	1.3	C4
26	1.3	1.4	1.4	1.3	1.4	0.3	1.4	1.4	1.4	0.3	C6
27	1.5	1.7	1.7	1.5	1.7	0.9	1.7	1.7	1.7	0.9	C6
28	1.7	1.9	1.7	1.3	1.7	1.4	2	2	1.9	1.3	C4
29	1.6	1.7	1.7	1.3	1.4	0.7	1.7	1.7	1.7	0.7	C6
30	1.4	1.9	1.7	1.5	2	1.6	2	2	1.8	1.4	C1
31	1.6	1.7	1.7	1.6	1.7	1.7	1.1	1.1	0.9	0.9	C9
32	1.3	2	2	1.9	2	2	1.7	2	1.8	1.3	C1

Data ke-	C1	C2	C3	C4	C5	C6	C7	C8	C9	Min	Kluster
33	1.4	1.7	1.7	1.7	1.7	1.7	0.8	1.7	1.4	0.8	C7
34	1.3	1.4	1.4	1.4	1.4	1.4	0.3	1.4	1	0.3	C7
35	1.3	1.7	1.7	1.7	1.7	1.7	1.1	1.4	0.9	0.9	C9
36	1.6	2	2	1.8	2	2	1.8	1.8	1.2	1.2	C9
37	1.6	1.7	1.7	1.7	1.7	1.7	1.5	1	1.3	1	C8
38	1.4	1.7	1.7	1.6	1.7	1.7	1.5	1	1.2	1	C8
39	1.5	1.7	1.7	1.6	1.7	1.7	1.4	1.5	0.9	0.9	C9
40	1.5	2	2	1.8	2	2	1.5	1.6	1	1	C9

Dari table 3.6 diatas didapatkan data sebagai berikut:

C1 = data ke-2, data ke-5, data ke-20, data ke-32

C2 = data ke-6, data ke-7, data ke-8, data ke-9, data ke-10,

C3 = data ke-13

C4 = data ke-16, data ke-17, data ke-19, data ke-23, data ke-25, data ke-28

C5 = data ke-21, data ke-22, data ke-24

C6 = data ke-26, data ke-27, data ke-29

C7 = data ke-1, data ke-4, data ke-33, data ke-34

C8 = data ke-11, data ke-14, data ke-15, data ke-37, data ke-38

C9 = data ke-3, data ke-12, data ke-31, data ke-35, data ke-36, data ke-39, data ke-40

Centroid iterasi ke-1 dan ke-2 belum sama maka dilakukan iterasi ke-3.

Dengan menggunakan rumus rata-rata dilakukan *update* terhadap *centroid*. Tabel hasil *update centroid* dapat dilihat pada tabel 3.7.

Tabel 3.7 hasil perhitungan jarak eucliden iterasi ke 3

Data ke-	C1	C2	C3	C4	C5	C6	C7	C8	C9	Min	Kluster
1	1.1	1.4	1.4	1.4	1.4	1.4	0.5	1	1	0.5	C7
2	1.3	2.2	2.2	2.1	2.2	2.2	2.1	2.2	1.9	1.3	C1
3	1.5	1.6	1.4	1.5	1.7	1.7	1.1	1.4	0.9	0.9	C9
4	1.8	2	2	1.9	1.8	2	0.9	1.7	1.4	0.9	C7
5	1.1	1.7	1.7	1.7	1.7	1.7	1.3	1.4	1.3	1.1	C1
6	2	0.8	2	2	2	2	2	2	2	0.8	C2
7	1.6	0.9	1.4	1.5	1.7	1.7	1.7	1.7	1.7	0.9	C2
8	1.4	0.7	1.4	1.4	1.4	1.4	1.4	1.4	1.4	0.7	C2
9	1.4	0.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	0.4	C2
10	1.4	0.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	0.4	C2
11	1.3	1.4	1.4	1.3	1.4	1.4	1.1	0.4	0.9	0.4	C8
12	1.6	1.7	1.7	1.7	1.7	1.7	1.1	1.3	0.9	0.9	C9
13	1.3	1.3	0.5	1.1	1.4	1.4	1.4	1.4	1.3	0.5	C3
14	1.6	1.7	1.7	1.6	1.7	1.7	1.5	0.9	1.3	0.9	C8
15	1.4	1.4	1.4	1.3	1.4	1.4	1.4	0.9	1.2	0.9	C8
16	1.6	1.7	1.7	1	1.1	1.6	1.7	1.7	1.7	1	C4
17	1.6	1.6	1.4	0.7	1.1	1.6	1.7	1.7	1.7	0.7	C4
18	1.3	1.3	0.5	1.1	1.4	1.4	1.4	1.4	1.3	0.5	C3
19	1.5	1.7	1.7	1	1.1	1.6	1.7	1.6	1.5	1	C4
20	1	1.4	1.4	1.2	1.4	1.4	1.4	1.3	1.2	1	C1
21	1.7	1.7	1.7	1.1	0.5	1.6	1.7	1.7	1.7	0.5	C5
22	1.7	1.7	1.7	1	0.5	1.6	1.6	1.7	1.7	0.5	C5
23	1.4	1.4	1.4	0.8	1	1.2	1.4	1.1	1.3	0.8	C4
24	1.7	1.7	1.7	1.3	0.5	1.6	1.6	1.7	1.7	0.5	C5
25	1.7	1.9	1.7	1.3	1.5	1.9	1.9	2	1.8	1.3	C4
26	1.3	1.4	1.4	1.3	1.4	0.3	1.4	1.4	1.4	0.3	C6
27	1.5	1.7	1.7	1.5	1.7	0.9	1.7	1.7	1.7	0.9	C6
28	1.7	1.9	1.7	1.3	1.7	1.4	2	2	1.9	1.3	C4
29	1.6	1.7	1.7	1.3	1.4	0.7	1.7	1.7	1.7	0.7	C6
30	1.4	1.9	1.7	1.5	2	1.6	2	2	1.8	1.4	C1
31	1.6	1.7	1.7	1.6	1.7	1.7	1.1	1.1	0.9	0.9	C9
32	1.3	2	2	1.9	2	2	1.7	2	1.8	1.3	C1

Data ke-	C1	C2	C3	C4	C5	C6	C7	C8	C9	Min	Kluster
33	1.4	1.7	1.7	1.7	1.7	1.7	0.8	1.7	1.4	0.8	C7
34	1.3	1.4	1.4	1.4	1.4	1.4	0.3	1.4	1	0.3	C7
35	1.3	1.7	1.7	1.7	1.7	1.7	1.1	1.4	0.9	0.9	C9
36	1.6	2	2	1.8	2	2	1.8	1.8	1.2	1.2	C9
37	1.6	1.7	1.7	1.7	1.7	1.7	1.5	1	1.3	1	C8
38	1.4	1.7	1.7	1.6	1.7	1.7	1.5	1	1.2	1	C8
39	1.5	1.7	1.7	1.6	1.7	1.7	1.4	1.5	0.9	0.9	C9
40	1.5	2	2	1.8	2	2	1.5	1.6	1	1	C9

Dari Tabel 3.7 didapatkan data sebagai berikut:

C1 = data ke-2, data ke-5, data ke-20, data ke-32

C2 = data ke-6, data ke-7, data ke-8, data ke-9, data ke-10,

C3 = data ke-13

C4 = data ke-16, data ke-17, data ke-19, data ke-23, data ke-25, data ke-28

C5 = data ke-21, data ke-22, data ke-24

C6 = data ke-26, data ke-27, data ke-29

C7 = data ke-1, data ke-4, data ke-33, data ke-34

C8 = data ke-11, data ke-14, data ke-15, data ke-37, data ke-38

C9 = data ke-3, data ke-12, data ke-31, data ke-35, data ke-36, data ke-39, data ke-40.

Tabel 3.7 menunjukkan iterasi ke-3 dengan pusat klaster yang telah diperbarui tidak mengalami perubahan (*konvergen*), maka dengan kata lain keanggotaan klaster pada iterasi ke-2 dan iterasi ke-3 adalah sama. Karena tidak mengalami perubahan anggota klaster, maka perhitungan tidak perlu dilakukan lagi dan proses iterasi telah berakhir.

(b) Naïve Bayes

Cara kerja dari proses perhitungan Naïve Bayes diawali dengan melakukan pengambilan data sampel atau data latih dari data yang telah melalui proses perhitungan dengan K-Means. Jika diasumsikan C9 adalah data 'mungkin' atau data yang belum terkategori. Dengan permisalan seperti pada Tabel 3.8.

Tabel 3.8 Pengelompokan Data Kluster

C1	scabies
C2	rhinitis
C3	otitis
C4	helminthiasis
C5	gastritis
C6	enteritis
C7	dermatophytosis
C8	dermatitis
C9	Mungkin (<i>uncategory</i>)

Maka dilakukan perubahan pengelompokan data, dari kluster n menjadi kelas-kelas seperti yang ditentukan sebelumnya. Setelah itu, dilakukan pemisahan data mungkin (*uncategory*). Data latih adalah data yang telah melalui perhitungan K-Means, sedangkan data mungkin adalah data yang akan diuji. Sehingga data mungkin (*uncategory*) yang akan digunakan dalam pengklasifikasian dapat dilihat pada Tabel 3.9.

Tabel 3.9 Data C9

Data ke	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29	G30	G31	G32	Kelas
3	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	dermatophytosis
12	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	scabies
31	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	dermatophytosis
35	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	scabies
36	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	rhinitis
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	rhinitis
40	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	rhinitis

Berikutnya adalah menghitung *prior probability* untuk masing-masing kelas dengan menggunakan rumus:

$$P(C_i) = \frac{\text{jumlah data pada masing-masing kelas}}{\text{total jumlah data}} \quad P(C_1) = \frac{5}{33} = 0,151515 \quad P(C_2) = \frac{5}{33} = 0,151515$$

Dengan cara perhitungan yang sama didapatkan hasil perhitungan *prior* seperti yang ditunjukkan pada Tabel 3.10.

Tabel 3.10 Hasil Perhitungan *Prior* tiap kelas

Prior	
Scabies	0.151515152
Rhinitis	0.151515152
Otitis	0.060606061
Helminthiasis	0.181818182
Gastritis	0.090909091
Enteritis	0.090909091
Dermatophytosis	0.121212121
Dermatitis	0.151515152

Selanjutnya adalah melakukan perhitungan *likelihood* untuk masing-masing data uji. Perhitungan *likelihood* dengan cara mengalikan nilai yang didapatkan dari perhitungan peluang G1 sampai G32 pada kelas Ci atau $P(G_i|C_j)$, dimana i merupakan banyaknya *feature* data dan j merupakan banyaknya kelas. Berikut adalah perhitungan *likelihood* dari data ke-3:

$$\begin{aligned}
 P(\text{Scabies}|X_3G1, X_3G2, \dots X_3G32) &= P(G_1|\text{Scabies}) * P(G_2|\text{Scabies}) * \dots * (G_{32}|\text{Scabies}) \\
 &= 0.8 * 0.2 * 0.2 * 0.6 * 0.6 * 0.2 * 0.4 * 1 * 1 * 0.8 * 0.8 * 1 * 1 * 1 * 0.8 * 0.8 * 0.8 * 1 * 0.8 * 1 * 1 * 1 * 0.8 * 1 * 1 * 1 * 1 * \\
 &\quad 1 * 1 * 1 * 1 * 0.8 \\
 &= 0.000154619
 \end{aligned}$$

Dengan cara yang sama dilakukan perhitungan untuk mencari nilai *likelihood* pada setiap kelas sehingga didapatkan hasil seperti yang ditunjukkan pada Tabel 3.11.

Tabel 3.11 Hasil Perhitungan *likelihood* data ke - 3

Likelihood	
Scabies	0.000154619
Rhinitis	0
Otitis	0
Helminthiasis	0
Gastritis	0
Enteritis	0
dermatophytosis	0
Dermatitis	0

Selanjutnya dilakukan perhitungan *posterior* dengan cara mengalikan *likelihood* dengan *prior*.

$$\text{Misalnya: } P(X_3 | \text{Scabies}) = P(\text{Scabies} | X_3G1, X_3G2, \dots, X_3G32) * P(\text{Scabies})$$

$$= 0.000154619 * 0.151515152 = 2.34271$$

Dengan cara yang sama dilakukan perhitungan posterior pada masing-masing kelas sehingga didapatkan hasil seperti yang ditunjukkan pada Tabel 3.12.

Tabel 3.12 Hasil Perhitungan posterior data ke - 3

Posterior	
Scabies	2.34271
Rhinitis	0
Otitis	0
Helminthiasis	0
Gastritis	0
Enteritis	0
dermatophytosis	0
Dermatitis	0
Posterior Tertinggi	2.34271
Hasil Klasifikasi	Scabies

Selanjutnya adalah dilakukan *likelihood* data ke-12, data ke-31, data ke-35, data ke-36, data ke-29, dan ke – 40 dengan cara yang sama. Berikut adalah perhitungan dari terakhir yaitu data ke-40.

$$\begin{aligned}
 P(\text{Scabies} | X_{40}G1, X_{40}G2, \dots X_{40}G32) &= P(G_1 | \text{Scabies}) * P(G_2 | \text{Scabies}) * \dots * (G_{32} | \text{Scabies}) \\
 &= 0.8 * 0.8 * 0.8 * 0.6 * 0.6 * 0.8 * 0.6 * 1 * 1 * 0.8 * 0.8 * 1 * 1 * 1 * 0.8 * 0.8 * 0.8 * 1 * 0.8 * 1 * 1 * 1 * 0.8 * 1 * 1 * 1 * 1 * \\
 &\quad 1 * 1 * 1 * 1 * 0.8 \\
 &= 0.014843407
 \end{aligned}$$

Dengan cara yang sama dilakukan perhitungan untuk mencari nilai *likelihood* pada setiap kelas sehingga didapatkan hasil seperti yang ditunjukkan pada Tabel 3.13.

Tabel 3.13 Hasil Perhitungan *likelihood* data ke - 40

Likelihood	
scabies	0.014843407
rhinitis	0.0589824
otitis	0.0625
helminthiasis	0.019845806
gastritis	0.043895748
enteritis	0.009754611
dermatophytosis	0.016685486
dermatitis	0.042467328

Selanjutnya dilakukan perhitungan posterior dengan cara mengalikan likelihood dengan prior.

Misalnya: $P(X_{40} | \text{Scabies}) = P(\text{Scabies} | X_{40}G1, X_{40}G2, \dots, X_{40}G32) * P(\text{Scabies})$

$$= 0.014843407 * 0.151515152 = 0.002249001$$

Dengan cara yang sama dilakukan perhitungan posterior pada masing-masing kelas sehingga didapatkan hasil seperti yang ditunjukkan pada Tabel 3.14.

Tabel 3.14 Hasil Perhitungan posterior data ke - 40

Posterior	
scabies	0.002249001
rhinitis	0.008936727
otitis	0.003787879
helminthiasis	0.003608328
gastritis	0.003990523
enteritis	0.000886783
dermatophytosis	0.002022483
dermatitis	0.006434444
Posterior Tertinggi	0.008936727
Hasil Klasifikasi	rhinitis

Dari hasil perhitungan *likelihood* mendapatkan nilai yang kemudian dibandingkan antar kelas dan diambil nilai tertinggi dari semua peluang kelas. Sehingga bisa didapatkan hasil dari akurasi keseluruhan dengan membandingkan antara kelas asli dengan kelas hasil *Naïve Bayes*. Kelas hasil *Naïve Bayes* disini adalah semuanya, dari kelas latih sampai kelas mungkin yang pada awalnya dijadikan kelas uji dan sudah memiliki kelas baru hasil dari perhitungan *likelihood*. Hasil pengujian dapat dilihat pada Tabel 3.15.

Tabel 3.15 Hasil pengujian

Data ke	Kelas Asli	Hasil Bayes	Akurasi
1	scabies	dermatophytosis	0
2	scabies	scabies	1
3	scabies	scabies	1
4	scabies	dermatophytosis	0
5	scabies	scabies	1
6	rhinitis	rhinitis	1
7	rhinitis	rhinitis	1
8	rhinitis	rhinitis	1
9	rhinitis	rhinitis	1
10	rhinitis	rhinitis	1
11	otitis	dermatitis	0
12	otitis	-	0
13	otitis	otitis	1
14	otitis	dermatitis	0
15	otitis	dermatitis	0
16	helminthiasis	helminthiasis	1
17	helminthiasis	helminthiasis	1
18	helminthiasis	otitis	0
19	helminthiasis	helminthiasis	1
20	helminthiasis	scabies	0
21	gastritis	gastritis	1
22	gastritis	gastritis	1
23	gastritis	helminthiasis	0
24	gastritis	gastritis	1
25	gastritis	helminthiasis	0
26	enteritis	enteritis	1
27	enteritis	enteritis	1
28	enteritis	helminthiasis	0
29	enteritis	enteritis	1

Data ke	Kelas Asli	Hasil Bayes	Akurasi
30	enteritis	scabies	0
31	dermatophytosis	scabies	0
32	dermatophytosis	scabies	0
33	dermatophytosis	dermatophytosis	1
34	dermatophytosis	dermatophytosis	1
35	dermatophytosis	dermatophytosis	1
36	dermatitis	dermatophytosis	0
37	dermatitis	dermatitis	1
38	dermatitis	dermatitis	1
39	dermatitis	rhinitis	0
40	dermatitis	rhinitis	0
Jumlah data benar			23
Akurasi			0.575

3.2.4.4 Perancangan Antarmuka Sistem

Pada perancangan antarmuka ini dijelaskan melalui desain antarmuka tiap halaman. Adapun tampilan antarmuka ditunjukkan pada Gambar 3.5 sampai Gambar 3.7.

(a) Rancangan Antarmuka Halaman Beranda

Klasifikasi Gejala Penyakit Kucing

Data **1** Browse

Detail Hitung

2

Akurasi K-Means vensional **3**

Akurasi K-Means Naive Bayes **4**

5 Submit

Gambar 3.5 Rancangan Antarmuka Halaman Beranda

Pada Gambar 3.5 yaitu tampilan awal sistem, berfungsi untuk menampilkan halaman awal. Pada halaman ini juga terdapat informasi sebagai berikut:

1. Menu Data berfungsi untuk meng-*input*-kan data.
2. Menu Detil Hitung menampilkan informasi sistem berupa Data Training, Pusat Awal, Distance Awal, Cluster, Pusat Akhir, Distance Akhir, Kluster, Hasil Bayes.
3. Menu Akurasi K-Measnvonseional menampilkan hasil akurasi K-Means Konvensional.
4. Menu Akurasi K-Means Naïve Bayes menampilkan hasil akurasi K-Means dan Naïve Bayes
5. Menu Submit untuk menambahkan data yang akan diproses

(b) Rancangan Antarmuka Halaman Data Training

Klasifikasi Gejala Penyakit Kucing

Data

Detil Hitung

Data

Gambar 3.6 Rancangan Antarmuka Halaman Data Training

Pada gambar 3.6 merupakan menu Data Training, menu ini berfungsi untuk menampilkan data yang telah di-*input*-kan.

(c) Rancangan Antarmuka Detil Perhitungan

Klasifikasi Gejala Penyakit Kucing

Data

Browse

K-Meansvensional

K-Means Naive Bayes

Detil Hitung

Akurasi K-Meansvensional

Akurasi K-Means Naive Bayes

Submit

Gambar 3.7 Rancangan Antarmuka Halaman Pengujian

Pada gambar 3.7 adalah menu Pengujian, yaitu menu yang berfungsi untuk menampilkan halaman pengujian. Halaman yang memuat hasil perhitungan dari data yang akan dimasukkan kedalam sistem oleh pengguna. Perhitungan menampilkan Data Training, Pusat Awal, Distance Awal, Cluster Awal, Pusat Awal, Distance Akhir, Hasil K-MEANS, Hasil Naïve Bayes.

3.2.5 Implementasi Sistem

Implementasi sistem adalah tahap membangun sistem klasifikasi yang sesuai dengan perancangan yang telah dibuat dan menerapkan hal yang telah didapatkan pada proses studi literatur. Aplikasi klasifikasi gejala penyakit kucing untuk mendeteksi penyakit kucing dengan menggunakan metode *K-Means* dan *Naïve Bayes*. Sistem dibangun menggunakan bahasa pemograman JAVA sebagai platform pengembangannya. Tahap implementasi yang dilakukan meliputi implementasi *interface*, basis data, dan algoritme. Keluaran yang dihasilkan dari implementasi sistem ini adalah diagnosis penyakit kucing.

3.2.6 Pengujian Sistem

Pengujian pada penelitian ini digunakan untuk mengetahui tingkat akurasi dan mengetahui apakah perangkat lunak yang dibuat telah memenuhi kebutuhan yang diinginkan. Pengujian sistem dilakukan dengan beberapa skenario uji coba. Pengujian tahap pertama yaitu data diklasifikasikan menggunakan algoritme *K-Means* dan *Naïve Bayes*. Selanjutnya nantinya *K-Means* akan menghasilkan sembilan klaster yaitu: scabies, gastritis, helminthiasis, rhinitis, dermatophytosis, dermatitis, enteritis, otitis dan mungkin. Untuk data yang belum masuk kategori (mungkin) akan dihitung menggunakan *Naïve Bayes*. Setelah itu tahap kedua, yaitu pengujian menggunakan data uji. Pada tahap ini pengujian menggunakan aturan yang telah terbentuk dari pengujian pada tahap pertama. Pengujian ini dilakukan untuk mengetahui ketetapan aturan yang dibuat dalam menganalisa data uji, sehingga pengklasifikasian terhadap data uji dapat dilakukan dengan benar. Selanjutnya pengujian dilakukan untuk mengukur tingkat akurasi algoritme yang digunakan pada proses klasifikasi terhadap data gejala penyakit kucing. Berikut pada Tabel 3.16 adalah perancangan uji coba sistem.

Tabel 3.16 Perancangan Uji Coba Sistem

Jumlah data latih (%)	Jumlah data uji (%)	Rata-rata Akurasi dari percobaan 1 sampai 5	Akurasi Tertinggi dari percobaan 1 sampai 5
60%	20%		
70%	20%		
80%	20%		

Pada Tabel 3.16. Pengujian menggunakan jumlah data latih yang berbeda dengan data uji. Data uji adalah tetap untuk melihat pengaruh jumlah data latih terhadap akurasi. Tahap awal perhitungan menggunakan sejumlah data latih sebesar 60% dari jumlah seluruh data. Aturan yang terbentuk dari data latih tersebut akan diujikan terhadap 20% data uji. Tahap selanjutnya, melakukan

perhitungan yang sama seperti dengan tahap awal, namun di tahap kedua menggunakan sejumlah data latih yang lebih besar yaitu 70% data latih dari jumlah seluruh data dan diujikan terhadap 20% data uji. Selanjutnya, tetap dilakukan perhitungan yang sama dengan sejumlah data latih 80% dari jumlah seluruh data dengan menggunakan jumlah data uji 20%.

Setelah melakukan perhitungan akan didapatkan nilai akurasi masing-masing perhitungan, kemudian dihitung rata-rata dari seluruh nilai akurasi dan didapatkan nilai akurasi tertinggi.

3.2.7 Kesimpulan

Pengambilan kesimpulan dilakukan setelah semua tahapan perancangan, implementasi dan pengujian metode yang diterapkan telah diterapkan selesai dilakukan. Kesimpulan diambil dari hasil pengujian dan analisis metode yang diterapkan. Selain itu penarikan kesimpulan dilakukan untuk menjawab rumusan permasalahan yang sudah ditetapkan sebelumnya. Tahap terakhir dari penulisan adalah saran yang bertujuan untuk memperbaiki kesalahan – kesalahan yang terjadi serta untuk memberikan masukan dan pertimbangan untuk penelitian yang selanjutnya.

BAB 4

IMPLEMENTASI DAN PEMBAHASAN

Pada bab ini akan membahas tentang implementasi berdasarkan metodologi dan perancangan yang telah ditetapkan. Pembahasan dalam tahap ini meliputi implementasi program dan implementasi antarmuka.

4.1 Lingkungan Implementasi

Pada lingkungan implementasi ialah menerapkan metode penelitian ke dalam sistem, perlu adanya beberapa faktor yang diperhatikan. Tujuannya untuk memenuhi kebutuhan dari sistem yang akan dikembangkan dan metode yang diimplementasikan. Lingkungan implementasi yang akan dijelaskan pada sub bab ini adalah lingkungan perangkat keras (*hardware*) dan lingkungan perangkat lunak (*software*) yang digunakan untuk membangun sistem klasifikasi gejala penyakit kucing menggunakan algoritme *K-Means* dan *Naïve Bayes*.

4.1.1 Lingkungan Implementasi Perangkat Keras

Perangkat keras yang digunakan dalam pembuatan sistem klasifikasi gejala penyakit kucing menggunakan algoritme *K-Means* dan *Naïve Bayes* adalah sebagai berikut :

1. Processor Intel® Celeron™ CPU N2840 @ 2.16GHz 2.16 GHz.
2. RAM 4.00 GB.
3. Harddisk 500 GB.
4. Monitor 14.1".
5. Keyboard.
6. Mouse.

4.1.2 Lingkungan Implementasi Perangkat Lunak

Perangkat lunak yang digunakan dalam pembuatan sistem klasifikasi gejala penyakit kucing menggunakan algoritme *K-Means* dan *Naïve Bayes* adalah sebagai berikut :

1. Sistem operasi Windows 10.
2. Bahasa pemrograman JAVA
3. Microsoft Excel 2010 sebagai pengolahan data.

4.2 Implementasi Program

4.2.1 Proses membuat *centroid* baru secara acak pada *K-Means*

Proses ini akan membuat *centroid* awal secara acak.

No	Source Code
1	<code>private double[][] pusatClusterawal() {</code>
2	<code> int check[] = new int[k];</code>
3	<code> pusat = new double[k][data[0].length];</code>
4	
5	
6	<code> for (int i = 0; i < check.length - 1; i++) {</code>
7	<code> check[i] = random.nextInt(data.length);</code>
8	
9	
10	<code> while</code>
	<code> (!kelas[i].equalsIgnoreCase(input.kelaslatih[check[i]]))</code>
	<code>) {</code>
11	<code> check[i] = random.nextInt(data.length);</code>
12	<code> }</code>
13	
14	
15	<code> }</code>
16	<code> check[check.length-1]= random.nextInt(data.length);</code>
17	<code> kelas[check.length-1]=input.kelaslatih[check[check.length-</code>
	<code>1]];</code>
18	<code> for (int i = 0; i < check.length; i++) {</code>
19	<code> for (int j = 0; j < data[0].length; j++) {</code>
20	<code> pusat[i][j] = data[check[i]][j];</code>
21	<code> }</code>
22	<code> }</code>
23	<code> return pusat;</code>

24	}
----	---

Source Code 4.1 Proses membuat *centroid* baru secara acak pada *K-Means*

Penjelasan Source Code 4.1 :

1. Baris ke-2 digunakan untuk mengisialisasi panjang *array* sejumlah kluster.
2. Baris ke-3 digunakan untuk membentuk sebuah *array* sepanjang jumlah fitur data untuk menampung titik *cluster* yang dibangkitkan.
3. Baris ke-6 dan ke-7 digunakan untuk me-random indeks sesuai jumlah data.
4. Baris ke-10 sampai baris ke-11 digunakan untuk mengecek indeks hasil random. Langkah ini diperlukan agar data yang dibangkitkan sebagai pusat *cluster* sesuai dengan masing-masing kelas. Artinya, pusat *cluster* dibangkitkan secara random tiap kelasnya.
5. Baris ke-16 sampai baris ke-20 digunakan untuk mengambil nilai fitur data yang sudah diambil indeksnya secara acak untuk kemudian dibangkitkan sebagai pusat *cluster*.
6. Baris ke-23 digunakan untuk mengembalikan pusat cluster yang telah dibangkitkan secara acak.

4.2.2 Proses menghitung jarak data ke *centroid* pada *K-Means*

Pada proses ini, data yang dimasukan akan dihitung jaraknya terhadap *centroid* yang ada dengan menggunakan rumus jarak *eulcedian*.

No	Source Code
1	<code>private double[][] distance(double pusat[][]) {</code>
2	
3	<code> distance = new double[data.length][k];</code>
4	<code> for (int i = 0; i < distance.length; i++) {</code>
5	<code> for (int j = 0; j < distance[0].length; j++) {</code>
6	<code> double temp = 0;</code>
7	<code> for (int l = 0; l < data[0].length; l++) {</code>
8	<code> temp+=Math.pow(data[i][l]-pusat[j][l],2);</code>
9	<code> }</code>
10	<code> distance[i][j] = Math.sqrt(temp);</code>
11	
12	<code> }</code>
13	<code> }</code>

14	}
15	return distance;
16	
17	}

Source Code 4.2 Proses menghitung jarak data ke *centroid* pada *K-Means*

Penjelasan Source Code 4.2 :

1. Baris ke-3 digunakan untuk menginstansiasi *array distance* dengan ukuran panjang `data.length` (banyak data) dan lebar `k` (banyak kelas).
2. Baris ke-4 digunakan untuk melakukan perulangan sebanyak jumlah data.
3. Baris ke-5 digunakan untuk melakukan perulangan sebanyak jumlah fitur.
4. Baris ke-7 digunakan untuk melakukan perulangan untuk menghitung *eucliden*.
5. Baris ke-8 digunakan untuk menghitung selisih jumlah antara data dengan pusat *cluster*.
6. Baris ke-10 digunakan untuk menghitung nilai jarak data ke-1 terhadap kelas `j`.
7. Baris ke-15 digunakan untuk mengembalikan nilai jarak.

4.2.3 Proses meng-update (memperbaharui) *centroid*

Pada proses ini melakukan *update centroid*. Untuk mendapatkan *centroid* baru dilakukan dengan menghitung rata-rata anggota *cluster* dari setiap *cluster*.

No	Source Code
1	<code>private double[][] pusatBaru(String cluster[], int count[]) {</code>
2	
3	
4	
5	<code>pusat=new double[pusat.length][pusat[0].length];</code>
6	<code>for (int i = 0; i < pusat.length; i++) {</code>
7	<code>for (int k = 0; k < data.length; k++) {</code>
8	<code>if(cluster[k].equalsIgnoreCase("C"+(i+ 1))) {</code>
9	<code>for (int j = 0; j < pusat[0].length; j++) {</code>
10	<code>pusat[i][j] += data[k][j];</code>

```

11         }
12     }
13 }
14 }
15 for (int i = 0; i < pusat.length; i++) {
16     for (int j = 0; j < pusat[0].length; j++) {
17         if (count[i] == 0) {
18             pusat[i][j] = 0;
19         } else {
20             pusat[i][j] = Math.abs(pusat[i][j] / count[i]);
21         }
22     }
23 }
24 }
25 }
26 return pusat;
27 }

```

Source Code 4.3 Proses meng-update centroid

Penjelasan *Source Code* 4.3:

1. Baris ke-5 digunakan untuk membuat *array* baru bernama *pusat* dengan ukuran panjang sejumlah *cluster* dan ukuran lebar sejumlah banyaknya fitur pada setiap pusat *cluster*.
2. Baris ke-6 digunakan untuk melakukan perulangan sebanyak jumlah *cluster*
3. Baris ke-7 digunakan untuk melakukan perulangan sebanyak jumlah data.
4. Baris ke-8 sampai baris ke-10 digunakan untuk menjumlahkan setiap fitur data yang memiliki *cluster* sama.
5. Baris ke-15 sampai baris ke-20 digunakan untuk membagi jumlah fitur yang memiliki *cluster* sama dengan jumlah data pada masing-masing *cluster*. Hasil pembagian jumlah fitur dengan jumlah data pada masing-masing *cluster* disimpan dalam variabel *pusat [i][j]* dan digunakan sebagai pusat *cluster* baru. Dan jika *cluster* tersebut tak memiliki anggota, maka pusat *cluster* yang dibangkitkan adalah 0.
6. Baris ke-26 digunakan untuk mengembalikan nilai pusat *cluster* baru.

4.2.4 Proses mendapatkan data latih untuk Naïve Bayes

No	Source Code
1	for (int i = 0; i < km.data.length; i++) {
2	if(km.clusterBaru[i].equalsIgnoreCase("C"+km.countC.length)) {
3	uji[check2] = km.data[i];
4	id[check2] = i;
5	check2++;
6	} else {
7	data[check1] = km.data[i];
8	kelas[check1] = km.clusterBaru[i];
9	check1++;
10	}
11	}

Source Code 4.4 Proses mendapatkan data latih untuk Naïve Bayes

Penjelasan Source Code 4.4:

1. Baris ke-1 digunakan untuk perulangan data c1 sampai c9.
2. Baris ke-2 digunakan untuk mengecek apakah *cluster* baru yang tersimpan sama dengan C km.countC.lenght atau c9.
3. Baris ke-3 digunakan untuk mengecek apakah hasilnya sama, jika sama maka akan disimpan dalam *array* uji untuk menyimpan data uji.
4. Baris ke-7 sampai baris ke-9, else jika tidak sesuai dengan pernyataan if atau dengan kata lain data yang sedang di *loop* memiliki nilai klaster c1 sampai c8, maka data akan disimpan dalam *array* data [check1] dan akan dipergunakan sebagai data latih.

4.2.5 Proses Menghitung *Prior Probability* pada Naïve Bayes

Proses ini menghitung *prior probability* dari tiap kelas.

No	Source Code
1	public void prior(Kmeans km) {
2	prior = new double[km.countC.length - 1];
3	jumlahData = 0;
4	for (int i = 0; i < km.countC.length - 1; i++) {
5	jumlahData += km.countC[i];

6	}
7	for (int i = 0; i < km.countC.length - 1; i++) {
8	prior[i] = km.countC[i] / jumlahData;
9	}
10	
11	}

Source Code 4.5 Proses Menghitung Prior Probability

Penjelasan Source Code 4.5:

1. Baris ke-2 digunakan untuk membuat sebuah array baru bernama prior dengan panjang sejumlah jumlah panjang *cluster* -1. Hal ini dikarenakan hanya akan ada 8 kelas dalam naive bayes, sedangkan pada proses sebelumnya terdapat 9 *cluster*.
2. Baris ke-3 hingga baris 6 digunakan untuk menghitung seluruh data yang sudah memiliki kelas.
3. Baris ke-7 digunakan untuk melakukan perulangan sebanyak jumlah kelas.
4. Baris ke-8 digunakan untuk mengitung *prior* ke *i* dengan cara membagi jumlah anggota kelas ke *i* dengan jumlah seluruh data.

4.2.6 Proses perhitungan Likelihood dan Posterior

1. Perhitungan Likelihood

No	Source Code
1	public double[][] likelihood(double uji[]) {
2	likelihood = new double[prior.length][data[0].length];
3	for (int i = 0; i < likelihood.length; i++) {
4	int check = 0;
5	for (int k = 0; k < data[0].length; k++) {
6	if (uji[k] == 1) {
7	likelihood[i][k] = priorAttribute[i][check];
8	
9	} else {
10	likelihood[i][k] = priorAttribute[i][check + 1];
11	
12	}
13	
14	check += 2;

15	
16	}
17	
18	}
19	return likelihood;
20	}

Source Code 4.6 Proses menghitung *likelihood* pada Naïve Bayes

Penjelasan *Source Code 4.6*:

1. Baris 2 digunakan untuk membuat *array* dengan ukuran panjang sejumlah kelas dan lebar sejumlah fitur data
2. Baris ke-4 digunakan untuk melakukan perulangan sebanyak kelas data.
3. Baris 3 hingga 18 digunakan untuk menemukan nilai *likelihood* tiap fitur data. Mulanya, variabel *Check* diinisialisasi = 0. Kemudian dilakukan perulangan sebanyak jumlah feature seperti yang nampak pada baris 5. Jika nilai feature = 1, maka nilai *likelihood* yang digunakan adalah *priorAttribute[i][check]*. Nmaun, jika nilai feature = 0, maka nilai *likelihood* yang digunakan adalah *priorAttribute[i][check+1]*. Setelah ditemukan nilai *likelihood* hingga indeks ke *[i][jumlah feature]*, maka *pointe i* akan bergerak ke indeks selanjutnya, yang mana *i* merupakan indeks yang menyatakan jumlah kelas.
4. Baris 19 digunakan untuk mengembalikan nilai *likelihood*.

2. Menghitung *Posterior*

No	Source Code
1	<code>public double[] posterior(double likelihood[][]) {</code>
2	<code> posterior = new double[prior.length];</code>
3	<code> for (int i = 0; i < posterior.length; i++) {</code>
4	<code> posterior[i] = 1;</code>
5	<code> for (int j = 0; j < likelihood.length; j++) {</code>
6	<code> posterior[i]=posterior[i] * likelihood[i][j];</code>
7	
8	<code> }</code>
9	
10	<code> }</code>
11	<code> return posterior;</code>
12	<code>}</code>

Source Code 4.7 Menghitung *Posterior*

Penjelasan *Source Code* 4.7:

1. Baris ke-2 digunakan untuk membuat *array* baru untuk menampung *posterior*.
2. Baris ke-3 digunakan untuk melakukan perulangan sebanyak jumlah *posterior*
3. Baris ke-5 digunakan untuk melakukan perulangan sebanyak jumlah *likelihood*.
4. Baris ke-6 digunakan untuk mengalikan *likelihood* dengan *prior* dan menyimpannya dalam *array posterior*.
5. Baris ke-11 digunakan untuk mengembalikan nilai *posterior*.

4.2.7 Proses Menghitung Akurasi

Pada proses ini untuk melakukan perhitungan akurasi.

No	Source Code
1	<code>akurasi = 0;</code>
2	
3	<code>for (int i = 0; i < all.length; i++) {</code>
4	
5	
6	
7	<code>if (all[i].equalsIgnoreCase(kmeans.input.kelaslatih[i])) {</code>
8	<code>akurasi++;</code>
9	<code>}</code>
10	<code>}</code>
11	<code>akurasi = akurasi / all.length;</code>
12	
13	
14	<code>}</code>

Source Code 4.8 Proses menghitung akurasi

Penjelasan *Source Code* 4.8:

1. Baris ke-3 digunakan untuk melakukan iterasi sebanyak jumlah data.
2. Baris 7 hingga baris 8 digunakan untuk menghitung jumlah data yang berhasil diklasifikasikan sistem dengan benar.

- Baris ke 11 digunakan untuk menghitung akurasi dengan cara membagi jumlah data yang benar dengan jumlah seluruh data.

4.2.8 Proses menentukan klas hasil uji

Pada proses ini akan menentukan kelas hasil uji berdasarkan label *cluster*.

No	Source Code
1	<code>public void hasil(int index, double posterior[]) {</code>
2	<code> hasil[index] = 0;</code>
3	<code> hasil1[index] = "";</code>
4	<code> for (int i = 0; i < posterior.length; i++) {</code>
5	<code> hasil[index] = Math.max(hasil[index], posterior[i]);</code>
6	<code> if (hasil[index] == 0) {</code>
7	<code> hasil1[index] = "-";</code>
8	<code> } else if (hasil[index] == posterior[i]) {</code>
9	<code> hasil1[index] = "C" + (i + 1);</code>
10	<code> }</code>
11	<code> }</code>
12	<code>}</code>
13	
14	
15	
16	
17	
18	<code>for (int i = 0; i < id.length; i++) {</code>
19	<code> kmeans.clusterBaru[id[i]] = hasil1[i];</code>
20	<code>}</code>
21	<code>for (int i = 0; i < kmeans.kelas.length - 1; i++) {</code>
22	<code> for (int j = 0; j < all.length; j++) {</code>
23	<code> if(kmeans.clusterBaru[j].contentEquals("C" + (i + 1))) {</code>
24	<code> all[j] = kmeans.kelas[i];</code>
25	<code> } else if(all[j]==null) {</code>
26	<code> all[j]="-";</code>
27	

Source Code 4.9 Proses Menentukan klas hasil uji

Penjelasan *Source Code* 4.9:

- Baris ke-4 dan 5 digunakan untuk mencari nilai *posterior* terbesar.
- Baris 6 hingga 10 digunakan untuk menentukan cluster berdasarkan nilai *posterior* terbesar. Jika *posterior* terbesar = 0, maka *cluster* = '-' atau tidak

- terdeteksi. Namun, jika *posterior* maksimal bukan 0, maka *cluster* = *cluster* ke $i+1$. Dimana i merupakan indeks yang memiliki *posterior* terbesar.
- Baris 18 dan 19 digunakan untuk menyimpan hasil klasifikasi *naive bayes* yang masih dilabeli dengan *cluster*.
 - Baris 21 digunakan untuk melakukan perulangan sejumlah kelas.
 - Baris 22 digunakan untuk melakukan perulangan sebanyak data.
 - Baris ke-24 hingga 27 digunakan untuk menentukan kelas hasil uji berdasarkan label *cluster*.

4.3 Implementasi Antarmuka

4.3.1 Antarmuka Utama

Gambar 4.1 Antarmuka Halaman Beranda

Pada Gambar 4.1 adalah menampilkan halaman awal sistem. Pada halaman ini juga terdapat informasi sebagai berikut:

- Menu Data berfungsi untuk meng-*input*-kan data.
- Memasukkan jumlah data latih (%)
- Menu *Submit* untuk menambahkan data yang akan diproses

4. Menu Detil Hitung menampilkan informasi sistem berupa Data Training, Pusat Awal, Distance Awal, Cluster Awal, Pusat Akhir, Distance Akhir, Hasil K-Means, Hasil Naïve Bayes.
5. Menu Akurasi K-Means untuk menampilkan hasil dari akurasi K-Means Konvensional.
6. Menu Akurasi K-Means Naïve Bayes untuk menampilkan hasil dari akurasi K-Means Naïve Bayes.
7. Menu *Submit* untuk menambahkan data yang akan diproses.

4.3.2 Antarmuka Aplikasi

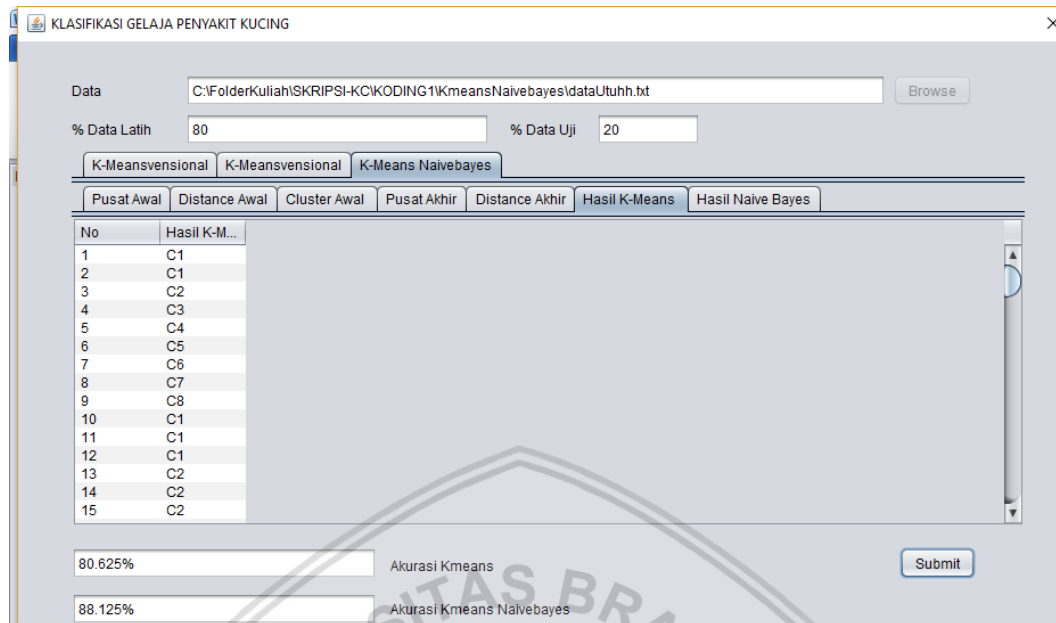
4.3.2.1 Proses Antarmuka Data *Training*

No	g1	g2	g3	g4	g5	g6	g7	g8	g9	g10
1	TIDAK	ADA	ADA	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK
2	ADA	TIDAK	ADA	ADA	ADA	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK
3	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK
4	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	ADA	ADA	TIDAK	TIDAK
5	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK
6	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	ADA	TIDAK	TIDAK	ADA	ADA
7	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK
8	TIDAK	TIDAK	TIDAK	TIDAK	ADA	TIDAK	ADA	TIDAK	TIDAK	TIDAK
9	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	ADA	ADA	TIDAK	TIDAK	TIDAK
10	ADA	TIDAK	ADA	TIDAK	TIDAK	ADA	ADA	TIDAK	TIDAK	TIDAK
11	TIDAK	ADA	ADA	TIDAK	TIDAK	ADA	ADA	TIDAK	TIDAK	TIDAK
12	TIDAK	TIDAK	ADA	ADA	ADA	TIDAK	ADA	TIDAK	TIDAK	TIDAK
13	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK
14	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK

Gambar 4.2 Antarmuka Data *Training*

Pada Gambar 4.2 merupakan menu Data Training, menu ini berfungsi untuk menampilkan data yang telah di-*input*-kan.

4.3.2.2 Proses Antarmuka Hasil K-MEANS



KLASIFIKASI GELAJA PENYAKIT KUCING

Data: C:\FolderKuliah\SKRIPSI-KC\KODING1\KmeansNaivebayes\dataUtuh.txt

% Data Latih: 80 % Data Uji: 20

K-Meansvensional K-Meansvensional K-Means Naivebayes

Pusat Awal Distance Awal Cluster Awal Pusat Akhir Distance Akhir Hasil K-Means Hasil Naive Bayes

No	Hasil K-M...
1	C1
2	C1
3	C2
4	C3
5	C4
6	C5
7	C6
8	C7
9	C8
10	C1
11	C1
12	C1
13	C2
14	C2
15	C2

80.625% Akurasi Kmeans

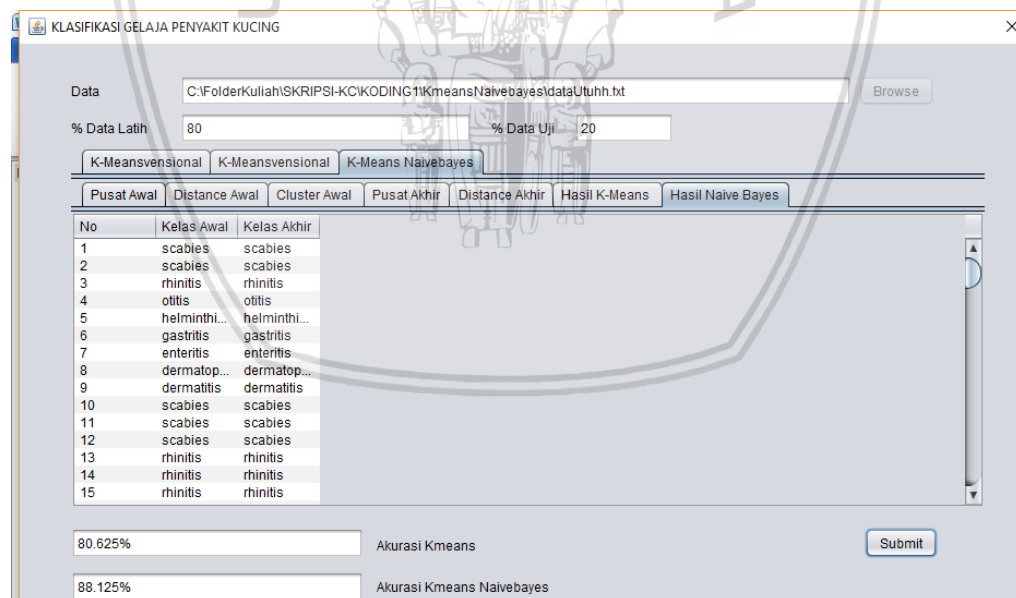
88.125% Akurasi Kmeans Naivebayes

Submit

Gambar 4.3 Antarmuka Hasil K-MEANS

Pada Gambar 4.3 menampilkan hasil perhitungan dari K-MEANS.

4.3.2.3 Proses Antarmuka Hasil Naïve Bayes



KLASIFIKASI GELAJA PENYAKIT KUCING

Data: C:\FolderKuliah\SKRIPSI-KC\KODING1\KmeansNaivebayes\dataUtuh.txt

% Data Latih: 80 % Data Uji: 20

K-Meansvensional K-Meansvensional K-Means Naivebayes

Pusat Awal Distance Awal Cluster Awal Pusat Akhir Distance Akhir Hasil K-Means Hasil Naive Bayes

No	Kelas Awal	Kelas Akhir
1	scabies	scabies
2	scabies	scabies
3	rhinitis	rhinitis
4	otitis	otitis
5	helminthi...	helminthi...
6	gastritis	gastritis
7	enteritis	enteritis
8	dermatop...	dermatop...
9	dermatitis	dermatitis
10	scabies	scabies
11	scabies	scabies
12	scabies	scabies
13	rhinitis	rhinitis
14	rhinitis	rhinitis
15	rhinitis	rhinitis

80.625% Akurasi Kmeans

88.125% Akurasi Kmeans Naivebayes

Submit

Gambar 4.4 Antarmuka Hasil Naïve Bayes

Gambar 4.4 menunjukkan hasil dari perhitungan Naive Bayes. Dimana ditampilkan kelas awal yang merupakan data awal sesuai kelas data. Kelas akhir merupakan hasil akhir dari *Naïve Bayes*.

BAB 5 PENGUJIAN DAN ANALISIS

Pada bab ini akan membahas pengujian dan analisis terhadap sistem yang telah diimplementasikan. Analisis dilakukan untuk melihat kekurangan-kekurangan pada sistem yang telah dibangun sehingga bisa dijadikan masukan untuk perbaikan maupun pengembangan terhadap sistem tersebut.

5.1 Pengujian Sistem

Pengujian sistem adalah hal terpenting yang bertujuan untuk menemukan kesalahan dan kekurangan pada sistem yang diuji. Pengujian dilakukan untuk mengetahui apakah sistem yang dibuat sudah memenuhi kriteria yang sesuai dengan tujuan perancangan sistem atau belum.

5.1.1 Penentuan Data Latih dan Data Uji

Data yang digunakan pada penelitian ini adalah 200 dataset yang berisikan gejala penyakit kucing dan disimpan dalam format txt. Pengujian untuk metode K-Means Konvensional dan metode gabungan K-Means dan Naïve Bayes menggunakan data dan prosentase dataset yang tetap dalam setiap kali percobaan. Pada penelitian ini, dataset diambil sebanyak 60%, 70% dan 80% data latih dan jumlah data uji yang sama yaitu 20%, sesuai dengan rancangan pengujian yang ditunjukkan pada Tabel 3.16.

5.1.2 Perbandingan Hasil Akurasi Antara Metode Gabungan K-Means dan Naïve Bayes (KMNB) Dengan Metode K-Means Konvensional.

Perbandingan hasil akurasi dari metode gabungan K-Means konvensional dengan metode K-Means dan Naïve Bayes bertujuan untuk mengetahui metode manakah yang menghasilkan akurasi tertinggi.

Pengujian terhadap metode K-Means konvensional dan metode K-Means Naïve Bayes akan dilakukan sebanyak 5 kali percobaan. Hasil akurasi akhir yang digunakan untuk mengukur kinerja sistem adalah hasil rata-rata dari akurasi 5 kali percobaan yang telah dilakukan.

5.1.2.1 Hasil Uji Coba Terhadap K-Means Konvensional

Akurasi dari metode K-Means konvensional dapat dilihat pada Tabel 5.1 sampai Tabel 5.3.

Tabel 5.1 Akurasi K-Means Konvensional dengan Data Latih 80% Data Uji 20%

	Percobaan 1	Percobaan 2	Percobaan 3	Percobaan 4	Percobaan 5
Akurasi	65,625%	68,125%	75,625%	65,25%	67,5%
Rata-rata uji	68,425%				

Tabel 5.1 menunjukkan sistem yang dibangun menggunakan K-Means konvensional dengan data latih 80% dari dataset dan 20% data uji dapat menghasilkan akurasi hingga 75,625% . Sedangkan setelah dilakukan 5 kali percobaan, rata-rata akurasi uji dengan data latih 80% dari dataset adalah 68,425%.

Tabel 5.2 Akurasi K-Means Konvensional dengan Data Latih 70% Data Uji 20%

	Percobaan 1	Percobaan 2	Percobaan 3	Percobaan 4	Percobaan 5
Akurasi	62,142%	68,571%	71,428%	72,857%	79,285%
Rata-rata uji	70,8566%				

Tabel 5.2 menunjukkan sistem yang dibangun menggunakan K-Means konvensional dengan data latih 70% dari dataset dapat menghasilkan akurasi hingga 79,285% . Sedangkan setelah dilakukan 5 kali percobaan, rata-rata akurasi uji dengan data latih 70% dari dataset adalah 70,8566%.

Tabel 5.3 Akurasi K-Means Konvensional dengan Data 60% Data Uji 20%

	Percobaan 1	Percobaan 2	Percobaan 3	Percobaan 4	Percobaan 5
Akurasi	66,06%	63,33%	75,83%	72,5%	79,16%
Rata-rata uji	71,379%				

Tabel 5.3 menunjukkan sistem yang dibangun menggunakan K-Means konvensional dengan data latih 60% dari dataset dapat menghasilkan akurasi hingga 79,16%. Sedangkan setelah dilakukan 5 kali percobaan, rata-rata akurasi uji dengan data latih 60% dari dataset adalah 71,379%.

5.1.2.2 Hasil Uji Coba Terhadap Metode Gabungan K-Means dan Naïve Bayes

Percobaan dilakukan dengan 5 kali percobaan untuk setiap data uji. Dalam setiap percobaan akan dicatat hasil dari akurasi dan rata-rata data uji. Berdasarkan perancangan pengujian pada Tabel 3.1 data yang latih yang akan digunakan adalah 60%, 70%, dan 80%, dengan jumlah data uji yang sama yaitu 20%. Hasil pengujian metode gabungan K-Means dan Naïve Bayes dapat dilihat pada Tabel 5.4 sampai Tabel 5.6.

Tabel 5.4 Akurasi K-Means Naïve Bayes dengan Data Latih 80% Data Uji 20%

	Percobaan 1	Percobaan 2	Percobaan 3	Percobaan 4	Percobaan 5
Akurasi	85,625%	89,375%	90,0%	90,625%	94,375%
Rata-rata uji	90%				

Tabel 5.4 menunjukkan sistem yang dibangun menggunakan K-Means Naïve Bayes dengan data latih 80% dari dataset dan 20 % data uji dapat menghasilkan akurasi hingga 94%. Sedangkan setelah dilakukan 5 kali percobaan, rata-rata akurasi uji adalah 90%.

Tabel 5.5 Akurasi K-Means Naïve Bayes dengan Data Latih 70% Data Uji 20%

	Percobaan 1	Percobaan 2	Percobaan 3	Percobaan 4	Percobaan 5
Akurasi	86,428%	89,285%	90,0%	91,428%	92,142%
Rata-rata uji	89,8566%				

Tabel 5.5 menunjukkan sistem yang dibangun menggunakan K-Means Naïve Bayes dengan data latih 70% dari dataset dan 20 % data uji dapat menghasilkan akurasi hingga 92,142% . Sedangkan setelah dilakukan 5 kali percobaan, rata-rata akurasi uji adalah 89,8566%.

Tabel 5.6 Akurasi K-Means Naïve Bayes dengan Data Latih 60% Data Uji 20%

	Percobaan 1	Percobaan 2	Percobaan 3	Percobaan 4	Percobaan 5
Akurasi	80,0%	88,33%	89,16%	90,83%	91,66%
Rata-rata uji	87,996%				

Tabel 5.6 menunjukkan sistem yang dibangun menggunakan K-Means Naïve Bayes dengan data 60% dari dataset dan 20 % data uji dapat menghasilkan akurasi hingga 91,66%. Sedangkan setelah dilakukan 5 kali percobaan, rata-rata akurasi uji adalah 87,966%.

5.2 Analisis Hasil Pengujian Metode K-Means dan Naïve Bayes (KMNB) dan Metode K-Means Konvensional

Pada tahap ini akan dilakukan analisis dari pengujian yang sudah dilakukan dengan metode *K-Means* dan *Naïve Bayes* untuk klasifikasi data gejala penyakit pada kucing. Analisis dilakukan untuk mengetahui pengaruh jumlah perbandingan data terhadap akurasi. Analisis hasil juga bertujuan untuk mengetahui hasil implementasi algoritme *K-Means* dan *Naïve Bayes* untuk mendeteksi gejala penyakit pada kucing. Pengujian Metode K-Means dan Naïve Bayes (KMNB) dan Metode K-Means konvensional dapat dilihat pada Tabel 5.7.

Tabel 5.7 Hasil Pengujian K-Means Naïve Bayes (KMNB) dan K-Means Konvensional

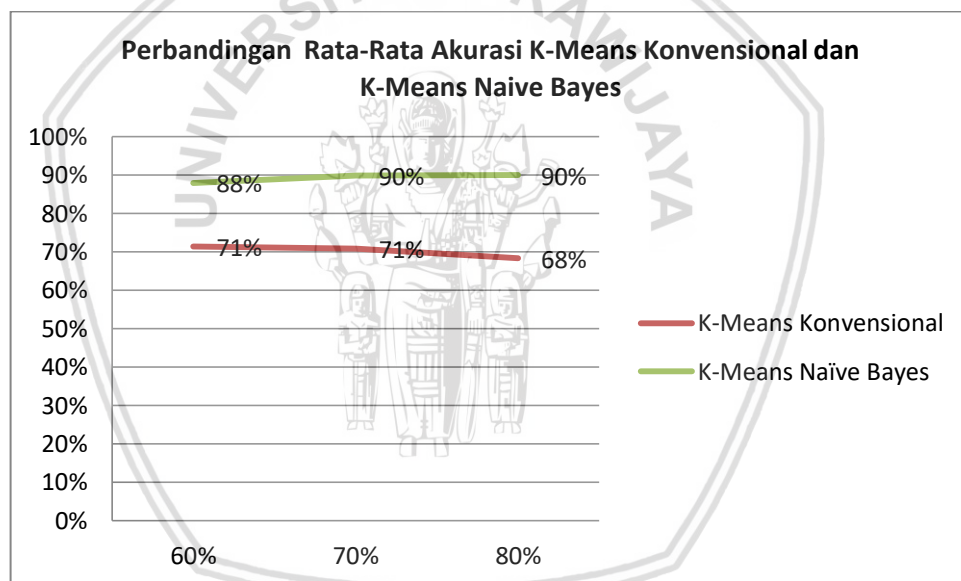
Jumlah Data (%)		Rata-rata Akurasi K-Means dan Naïve Bayes	Rata-rata Akurasi K-Means Konvensional
Data Latih	Data Uji		
80%	20%	90%	68, 425%
70%	20%	89, 8566%	70, 8566%
60%	20%	87, 996%	71, 379%

Pada Tabel 5.7 diketahui jika akurasi sistem yang dibangun menggunakan K-Means Naïve Bayes memiliki rata-rata akurasi terbaik ketika sistem dijalankan menggunakan data latih sebanyak 80% dari dataset dan data uji sebanyak 20%. Rata-rata akurasi terbesar yang bisa diraih sistem adalah 90%.

Akurasi tertinggi yang dicapai pada data latih sebanyak 80% dari dataset menunjukkan jika semakin banyak data latih yang digunakan akan semakin besar akurasi sistem. Hal ini berkaitan dengan proses pelatihan pada algoritme Naïve Bayes. Semakin banyak data yang digunakan untuk pelatihan, maka semakin besar akurasi sistem yang dihasilkan.

Pada pengujian K-Means konvensional dapat diketahui jika rata-rata akurasi terbaik didapatkan ketika sistem dijalankan menggunakan 60% data input. Rata-rata akurasi terbesar yang bisa diraih sistem adalah 71, 379%. Akurasi tertinggi yang dicapai pada data input sebanyak 60% dari dataset menunjukkan jika semakin sedikit data yang dikelompokkan maka semakin tinggi nilai akurasi. Hal ini dipengaruhi oleh adanya kemungkinan data yang memiliki gejala serupa tapi memiliki kelas berbeda semakin besar jika dataset yang digunakan semakin banyak.

Perbandingan tingkat rata-rata akurasi antara metode K-Means konvensional dan gabungan K-Means Naïve Bayes (KMNB) dapat dilihat pada Gambar 5.1.



Gambar 5.1 Grafik Perbandingan rata-rata akurasi sistem

Dari Gambar 5.1 diketahui jika metode gabungan K-Means dan Naïve Bayes dapat menghasilkan akurasi lebih baik dibandingkan K-Means konvensional. Perbedaan tingkat akurasi ini dikarenakan data yang tak terkluster pada K-Means konvensional dapat diminimalisir dengan adanya penambahan metode Naïve Bayes untuk proses klasifikasi.

BAB 6 PENUTUP

6.1 Kesimpulan

Berdasarkan hasil penelitian implementasi metode K-Means dan Naïve Bayes untuk klasifikasi gejala penyakit kucing, dapat diambil kesimpulan:

1. Metode K-Means dan Naïve Bayes dapat diimplementasikan untuk mengklasifikasi gejala penyakit kucing. Dalam penelitian ini, algoritme K-Means akan membagi dataset menjadi 9 cluster, yaitu (Scabies, Rhinitis, Ototis, Helmanthiasis, Dermatitis, Dermathopytosis, Gastritis, Enteritis dan Mungkin). Data Mungkin kemudian akan di proses menggunakan algoritme Naïve Bayes sehingga hasil keluaran akhir sistem hanya terdiri dari 8 kelas data (Scabies, Rhinitis, Ototis, Helmanthiasis, Dermatitis, Dermathopytosis, Gastritis dan Enteritis).
2. Pengujian dengan variasi data latih 60%, 70% dan 80% dari dataset menunjukkan jika metode gabungan K-Means dan Naïve Bayes (KMNB) mampu menghasilkan akurasi lebih baik dibandingkan K-Means konvensional. K-Means konvensional menghasilkan rata-rata akurasi terbaik sebesar 71,379%, dengan nilai rata-rata akurasi 68,425%, 70,8566%, dan 71,379%. Metode gabungan K-Means dan Naïve Bayes mampu menghasilkan rata-rata akurasi terbaik sebesar 90%. Nilai rata-rata akurasi KMNB adalah 87,996%, 89,8566%, dan 90%. Hal ini menunjukkan jika semakin banyak data latih yang digunakan maka semakin besar akurasi yang dihasilkan sistem.

6.2 Saran

Saran yang diberikan sebagai bahan pengembangan penelitian lanjutan adalah :

1. Penerapan Metode gabungan K-Means dan Naïve Bayes (KMNB) dengan permasalahan lain untuk mengukur kinerja sistem terhadap permasalahan serupa.

2. Penggunaan metode K-Means yang digabungkan dengan metode Naïve Bayes menghasilkan akurasi yang cukup tinggi jika dibandingkan dengan metode K-Means Konvensional, maka tidak menutup kemungkinan jika metode K-Means bisa bekerja lebih baik jika digabungkan metode lainnya.



DAFTAR PUSTAKA

- Anggreyni. 2015. *Implementasi Metode Clustering untuk Klasifikasi Kanker Payudara Menggunakan Algoritme Naïve Bayes dan K-Means*. Fakultas Ilmu Komputer, Universitas Brawijaya, Malang.
- Anonim. 2011. Dermatitis pada Anjing. *World of Vet*. [Online] Tersedia di worldofvet.blogspot.com/2011/11/dermatitis-pada-anjing.html. [Diakses Pada tanggal 5 Oktober 2016]
- Defianti Sofi. 2015. Integrasi Metode Klasifikasi Dan Clustering dalam Data Mining. Konferensi Nasional Informatika (KNIF). pp 39
- Giyanto, Heribertus. 2008. *Penerapan algoritme Clustering K-Means, K-Medoid, Gath Geva*. Tesis Tidak Terpublikasi. Yogyakarta: Universitas Gajah Mada.
- Gustiar Rendra. 2011. Helminthiasis. *VeterinerGustiar*. [Online] Tersedia di <https://veterinergustiar.wordpress.com>. [Diakses Pada tanggal 8 Oktober 2016]
- Meilani D B. 2014. APLIKASI DATA MINING UNTUK MENGHASILKAN POLA KELULUSAN SISWA DENGAN METODE NAÏVE BAYES. *JURNAL LINK*. pp. 3.
- Natalius. 2011. Metoda Naïve Bayes Classifier dan Penggunaannya pada Klasifikasi Dokumen. *Makalah I12092 Probabilitas dan Statistik*. p.1.
- Neno. 2016. Infeksi Telinga (Otitis) pada Kucing. *Kucingkita.com*. [Online] Tersedia <http://www.kucingkita.com/>. [Diakses Pada tanggal 9 Oktober 2016].
- Palguna David dkk. 2014. SISTEM PAKAR DIAGNOSIS PENYAKIT KULIT PADA KUCING MENGGUNAKAN METODE CERTAINTY FACTOR. *JSIKA*. pp. 75.
- Pandeeswari, at al. 2015. *K-Means Clustering and Naive Bayes Classifier For Categorization Of Diabetes Patient*. *International Journal of Innovative Science, Engineering & Technology*.
- Putra Alimansyah. 2015. Kasus Rhinitis pada Kucing. *Meet Vet*. [Online] Tersedia <http://meevet.blogspot.co.id/2012/12/kasus-rhinitis-pada-kucing.html>. . [Diakses Pada tanggal 5 Oktober 2016].

- Ratnawati, Dian Eka. 2018. *A Modified K-Means with Naïve Bayes (KMNB) Algorithm for Breast Cancer Classification*. Journal of Telecommunication, Electronic and Computer Engineerin.
- Saputra Thomas. 2015. Enteritis Pada Hewan. Ilmu Veteriner. [Online] Tersedia. <http://ilmuveteriner.com/enteritis-pada-hewan/> [Diakses Pada tanggal 7 Oktober 2016]
- Suwed, dkk. 2011. *Panduan Lengkap Kucing*. Bogor: Penebar Swadaya.
- Triakoso N. 2013. Small Animal Dematology. [Online] Tersedia <https://triakoso.wordpress.com/tag/dermatologi/> [Diakses Pada tanggal 5 Oktober 2016]
- Vhodzan Adzima. 2013. ISOLASI DAN IDENTIFIKASI KAPANG PENYEBAB DERMATOFITOSIS PADA ANJING DI KECAMATAN SYIAH KUALA BANDA ACEH. *Jurnal Medika Veterinaria*. pp.46
- Zaki Meira. 2014. DATA MINING AND ANALYSIS. *Fundamental Concepts and Algorithms*. p. 1.
- Z. Muda et al, W. Y. M. S. N. U., 2011. *A K-Means and Naive Bayes Learning Aproach for Better Intrusion Detection*.